REPORT RESUMES

ED 002 899 24

THE PERCEPTION OF MUSIC SYMBOLS IN MUSIC READING BY NORMAL CHILDREN AND BY CHILDREN GIFTED MUSICALLY.

BY- PETZOLD, ROBERT G.
WISCONSIN UNIV., MADISON, SCH. OF EDUC.
REPORT NUMBER CRP-554 PUB DATE 59
CONTRACT OEC-SAE-8167
EDRS PRICE MF-$0.27 HC-$5.52 138P.

DESCRIPTORS- GIFTED, *MUSIC TECHNIQUES, *MUSIC EDUCATION, *ELEMENTARY SCHOOL STUDENTS, COMPARATIVE TESTING, *TALENTED STUDENTS, GROUP TESTS, MUSIC ACTIVITIES, GRADE 4, GRADE 5, GRADE 6, MUSIC READING, MADISON, WISCONSIN

THIS STUDY WAS CONCERNED WITH IDENTIFYING THE DIFFERENCES BETWEEN CHILDREN OF AVERAGE MUSICAL ABILITY AND CHILDREN GIFTED MUSICALLY AS THEY ENGAGED IN CERTAIN MUSIC ACTIVITIES. DATA WERE OBTAINED FROM TESTS OF 227 MADISON, WISCONSIN, PUBLIC SCHOOL CHILDREN, GRADES 4-6. THE TAPE-RECORDING TEST WAS IN 2 PHASES--(1) 89 CHILDREN, SELECTED AT RANDOM AND UNDIFFERENTIATED IN TERMS OF MUSICAL ABILITY, WERE GIVEN 3 TRIALS TO ATTEMPT TO LEARN A SET OF 10 DIFFERENT TONAL CONFIGURATIONS, AND (2) 138 AVERAGE AND GIFTED CHILDREN WERE GIVEN 8 TRIALS TO LEARN 5 TONAL CONFIGURATIONS AND A SIMILAR NUMBER OF TRIALS TO LEARN A SONG USING THE SAME CONFIGURATIONS. THE RESULTS OF THE STUDY SHOWED NO SIGNIFICANT DIFFERENCES BETWEEN BOYS AND GIRLS IN THEIR ABILITY TO READ MUSIC OR BETWEEN GRADE LEVELS ON KNOWLEDGE OF MUSICAL NOTATION. HOWEVER, SIXTH-GRADE STUDENTS LEARNED THE TASKS AT A SUPERIOR RATE. MUSICALLY GIFTED STUDENTS READ MUSIC AT A RATE THREE OR FOUR TIMES FASTER. AURAL-ORAL IMITATION OF ITEMS WAS CONSIDERED EASIER THAN RESPONDING TO VISUAL STIMULI, BUT SUBJECTS WERE UNABLE TO PERCEIVE A GROUPING OF NOTES AS A WHOLE. THE STUDY TENDED TO SUPPORT THE POINT OF VIEW THAT A HIGHER LEVEL OF MUSIC READING COMPETENCE DEPENDS UPON PROVIDING CHILDREN WITH ACTIVITIES WHICH WILL ENABLE THEM TO UNDERSTAND AND USE THE CONCEPT UNDERLYING NOTATION. (GC)

ED002899

THE PERCEPTION OF MUSIC SYMBOLS IN MUSIC READING BY NORMAL CHILDREN AND BY CHILDREN GIFTED MUSICALLY

COOPERATIVE RESEARCH PROJECT #554

DR. ROBERT GEORGE PETZOLD

UNIVERSITY OF WISCONSIN

ED002819

Put the [illegible] by [illegible] report on the
3414
[illegible] 4, 766, 1051

PROJECT SAE 8167

THE PERCEPTION OF MUSIC SYMBOLS IN MUSIC READING BY NORMAL CHILDREN AND BY CHILDREN GIFTED MUSICALLY

Dr. Robert G. Petzold

ABSTRACT

This study was concerned with identifying the differences between children of average musical ability and children gifted musically as they engaged in certain music reading activities. To obtain the data needed for such comparisons, 227 fourth, fifth and sixth grade children from the Madison, Wisconsin, public elementary schools were tested. The testing was all individual with the responses tape-recorded for later analysis. The study was divided int two distinct parts: Phase I, in which 89 children, selected at random from the three grades and undifferentiated in terms of musical ability, were given three trials in which to attempt to learn a set of ten different tonal configurations; and Phase II, in which 138 average and gifted fourth and sixth grade children were given eight trials in which to learn five tonal configurations and a similar number of trials to learn a song using the same configurations. For Phase II these two tasks were presented on different days, with the order of the tasks reversed for half of the sample. The content of the tasks was based on an analysis of 326 songs from several of the basic song-texts, resulting in the identification of twenty common tonal configurations appearing in the songs. A total of 558 different configurations had been identified, with a total frequency count of 4575.

The fifteen-minute testing period for each of the tasks was not long enough for most subjects to learn the material but a great deal of significant information was obtained. The testing procedures proved workable and the data obtained is of considerable help in determining the nature of future investigations in this problem area of music reading.

The results of the study may be briefly summarized as follows:

1. No significant differences were found between boys and girls in terms of their ability to read music.
2. No significant differences were found between the grades when the scores used for comparison reflected the subject's knowledge about musical notation.

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

3. When the fourth and sixth grades were compared on the basis of the rate at which they learned the tasks the sixth grade was found to be significantly superior.
4. The musically gifted subjects consistently did better than average subjects at both grade levels. They knew more about reading music and learned the material at a rate three to four times faster than average subjects.
5. No significant differences, in terms of ability to read music, were found between subjects who had at least one or more years of instrumental training and subjects with no such training.
6. Learning the tonal configurations was not significantly influenced by prior practice on learning a song while learning the song showed prior practice in learning tonal configurations to be of considerable help.
7. The degree to which tonal configurations retained their identity from one context to another varied from group to group within and between grades. The inconsistent performance of the subjects in each of the groups from one task to another partially accounted for the lack of an overall consistent pattern of identification. In general, the sixth grade was better able to recognize the more complex configurations as they appeared in a different context while the fourth grade subjects showed greater recognition of the simpler configurations.
8. The subjects had considerable difficulty in accurately identifying tonal configurations within the context of the song because their responses were so strongly influenced by the nature of the surrounding musical material. The configuration was performed not as it actually appeared but as the subject thought it should appear in reference to what preceded the given tonal configuration in the song.
9. Analysis of the reading errors of tonal configurations showed that subjects do considerable guessing when reading and, while they perceive and respond to the outward shape of the configuration, they lack the ability to accurately identify the more subtle changes which differentiate between configurations which are similar but not identical. This is probably attributed to the fact that most of the music reading was a note-by-note process and the subjects were unable to perceive a grouping of notes as a whole.
10. Accurate aural perception of tonal configurations is usually accompanied by accurate visual perception of these same configurations although the

correlation between the two types of responses did not exceed .70 for any group.

The results of this study tend to support the point-of-view that a higher level of music reading competence depends upon providing children with activities which will enable them to understand and use the concepts underlying musical notation. They need a working knowledge of the "tools" of musical notation but will be unable to use these tools effectively unless greater emphasis is given to those activities which give the child an opportunity for independent performance, both group and individual. Such independent performance implies that the teacher should not always feel compelled to assist the class with her voice or with some type of instrumental accompaniment. Children in this study demonstrated that they did not read music successfully unless they thought about the notation and about the reading process. There is no incentive or need to think unless the reading activity is interesting, challenging, and relevant to other kinds of musical activities taking place within the school program.

University of Wisconsin

Final Report

June 30, 1959

Project SAE 8167 (January 1 to June 30, 1959)

THE PERCEPTION OF MUSIC SYMBOLS IN MUSIC READING BY NORMAL CHILDREN AND BY CHILDREN GIFTED MUSICALLY

Dr. Robert G. Petzold

School of Education
Madison, Wisconsin

INTRODUCTORY NOTE

This final report supplements and completes the brief progress report dated March 30, 1959. This project, SAE8167, was authorized for a six month period, January 1 through June 30, 1959. The testing of the 89 subjects in Phase I was completed on March 27, 1959, and the testing of the 138 subjects in Phase II was completed on May 28, 1959.

The several sections of the report contain all of the information considered significant to the study. Although this has been considered an exploratory study of the problem the writer believes the results of the study carry beyond the usual limits of such an exploratory study.

ACKNOWLEDGEMENT

"The research reported herein was performed pursuant to a contract with the United States Office of Education, Department of Health, Education, and Welfare."

CONTENTS

SECTION Page

I. INTRODUCTION 1

Introductory Statement
The Need for Additional Research
The Scope of the Study
The Purposes of the Study
The Procedures of the Study

II. SUMMARY OF RELATED INFORMATION 8

Purposes of the Music Program
Purposes of Music Reading
Research Literature

III. ANALYSIS OF SONG MATERIALS 18

IV. SELECTION OF THE SAMPLE 26

Preliminary Statement
Selection of the Population for the Study
Selection of the Sample for Phase I
Selection of the Sample for Phase II

V. PROCEDURES OF THE STUDY 35

Phase I
Phase II

VI. RESULTS OF THE STUDY, PHASE I 58

Visual Part of Phase I
Analysis of Reading Errors for the Visual Part of Phase I
Test Scores on the Aural Part of Phase I
Analysis of Singing Errors for the Visual Part of Phase I

VII. RESULTS OF THE STUDY, PHASE II 84

Subjects Tested
Test Results, Tonal Configurations
Test Results, Song
Test Results, Identity of Items
Influence of Instrumental Training on Test Scores
Relationship between Scores on Tonal Configurations, Teachers Ratings, and Kwalwasser Test

VIII. SUMMARY AND CONCLUSIONS 116

IX. BIBLIOGRAPHY 122

X. APPENDIX 124

TABLES

Page

1. Table Showing the Frequency Count of 558 Tonal Configurations 21

2. Table Listing the 22 Most Frequently Used Tonal Configurations as Identified in the Present Study and by Smiley 24

3. Table Showing the Frequency Arrangement of 25 Tonal Configurations having a Frequency Count of 20 to 39 . . 25

4. Table Showing the Relative Sizes of the 22 Elementary Schools in Madison, Wisconsin 28

5. Table Showing the Distribution of Subjects Participating in Phase I 28

6. Table Showing the Mean Average Scores and Standard Deviations of the Teacher Ratings 31

7. Table Summarizing the Scores Earned on the Kwalwasser Music Talent Test by 438 4th and 6th Grade Pupils . . . 31

8. Table Showing Pupil Enrollments in the 4th and 6th Grade for Each of the Three Schools 34

9. Table Showing the Representative Sample Initially Selected for the Testing Program of Phase II 34

10. Table Showing the Percentages of the Upper and Lower Groups making Correct Responses to the Visual Presentation of the 10 Items 47

11. Table Showing the Percentages of the Upper and Lower Groups making Correct Responses to the Aural Presentation of the 10 Items 47

12. Table Showing the Percentages of the Upper and Lower Groups in each Grade making Correct Responses to the Visual Presentation of the 10 Items 48

13. Table Showing the Initial Assignment of the 151 Subjects Participating in Phase II 51

14. Table Showing the Mean Scores for the Visual Part of Phase I for each Grade in Each School 59

15. Analysis of Variance between Grades for the Visual Part of Phase I 60

16. Analysis of Variance between Schools for the Visual Part of Phase I 61

Tables (Cont'd)

Page

17. Table Showing the Mean Scores for the 43 Boys and 46 Girls Tested in Phase I 62

18. Table Summarizing the Responses made by 89 Subjects to the Visual Part of Phase I 64a

19. Table Showing the Mean Scores for the Aural Part of Phase I 73

20. Analysis of Variance between Grades for the Aural Part of Phase I 75

21. Analysis of Variance between Schools for the Aural Part of Phase I 75

22. Table Summarizing the Responses made by 89 Subjects to the Aural Part of Phase I 78a

23. Table Summarizing the Responses made by 89 Subjects to both the Visual and Aural Parts of Phase I 79

24. Table Summarizing the Responses made by 89 Subjects to the Aural Part of Phase I (Combined Grades) 79

25. Table Showing the Distribution of the 138 Subjects Tested in Phase II 85

26. Table Summarizing the Age and I.Q. of the 138 Subjects in Phase II 85

27. Table Summarizing the Mean Number of Months of Instmumental Instruction for 53 Subjects in Phase II 87

28. Table Showing Per Cent of Sample Reporting Instrumental Instruction 87

29. Table Showing the "Per cent correct" Mean Scores for Two Equal Groups 89

30. Table Showing the Mean "Per cent Correct" Scores for Average and Gifted Subjects on the Tonal Configurations 91

31. Table Showing the Mean "Rate of Learning" Scores for Average and Gifted Subjects on the Tonal Configurations 91

32. Analysis of Variance between Schools for the Tonal Configurations of Phase II 96

33. Table Showing the Mean "Per cent Correct" Scores for Average and Gifted Subjects on the Song 99

34. Table Showing the Mean "Rate of Learning" Scores for Average and Gifted Subjects on the Song 99

35. Analysis of Variance Between Schools for the Song of Phase II . 101

FIGURES

Page

1. The Twenty Test Items Initially Selected for Phase I 37
2. The Ten-Item Test for Phase I 40
3. The Five-Item Test for Phase II 46
4. The Song for Phase II 50
5. Data Sheet, Phase II, Song 57

THE PERCEPTION OF MUSIC SYMBOLS IN MUSIC READING
BY NORMAL CHILDREN AND BY CHILDREN GIFTED MUSICALLY

(Project SAE 8167)

INTRODUCTION

Introductory Statement

The complex patterns of sounds which are called music may, subject to certain limitations, be indicated with a relatively high degree of accuracy through the use of an equally complex system of musical notation. This system of notation utilizes symbols which may be grouped into three general categories, each category being determined by the nature and the function of the symbol: (1) symbols which indicate the pitch of musical sounds; (2) symbols which indicate the temporal duration of musical sounds; and (3) symbols which indicate certain related aspects of musical interpretation such as tempo, dynamics, phrasing, and style of performance. Music reading becomes a highly complex process because it implies not only the recognition and interpretation of the symbols in each of these categories, but frequently requires a simultaneous response in terms of vocal or instrumental performance. The process is further complicated when the singer is called upon to read words which are placed on a level different than that of the notation, or when the instrumentalist is required to read more than a single line of notes, as is the case with piano or organ music.

In order to provide opportunities for all children to acquire a minimal level of music reading competence the music program of the school gives considerable emphasis to a variety of activities designed to facilitate the development of this skill. There is probably no single topic that has provoked more discussion and debate over a period of years than this problem of music reading. An exception might be the research and discussion which has centered around the problem of determining and measuring the nature of musical talent or native capacity. Unfortunately, there has been very little systematic research concerned with the several aspects of music reading although there has been considerable discussion regarding the following:

1. Adequate identification of the learning process as it applies to music reading.
2. Development of effective procedures for teaching music reading.
3. Designing effective instruments to measure and evaluate music reading competence.
4. Determining reasonable levels of music reading competence for children representing various degrees of musical ability.

The Need for Additional Research in Music Reading

The need for an investigation of this kind arises out of the lack of information concerning the nature of the music reading process as carried on by children of elementary school age. Many music educators recommend that music reading skills can be developed by learning a "vocabulary" of tonal patterns, or configurations, which can then be applied to new situations. These recommendations have been implemented in many of the basic music texts currently used in the elementary grades through actual identification of such configurations in the songs by means of brackets or taking them out of context as drill patterns. There are also many music teachers who utilize the "pattern" approach in their teaching, with frequent drill work by the students using flash cards, chalkboard work, games of all kinds and other such devices. The basic assumption underlying this approach is that such tonal configurations, when learned by the child, retain their visual identity in different contexts. While the professional literature shows almost complete agreement that music reading is a response to a group of notes rather than to single notes, there is less agreement as to the procedures to be followed. Is there a standard list of tonal configurations which should be learned apart from the song materials themselves, or should the tonal configurations to be learned be taken directly out of the song materials for study and then returned to their original context? The latter point-of-view suggests that there should be no standard list of such configurations and would seem to give emphasis to mastery of an identification process rather than mastery of standardized content.

The Scope of the Study

This study will deal primarily with reading those music symbols which indicate pitch relationships with other kinds of symbols to be considered in future investigations. Several terms may be used to identify either the process of reading music, or to identify a specific kind of music-reading activity. The professional literature shows these terms to be frequently interchanged, leading to ambiguities in application and confusion in interpretation. The commonly used terms, together with definitions, may be briefly summarized as follows:

1. Music reading is the process of reading and interpreting the many kinds of music symbols by converting them silently into tones having duration, pitch, intensity, and interpretative characteristics. At this level the process is an internal one which does not require a physical response to the symbols in order for the meaning of the symbols to become apparent. This represents a high level of musical

development in terms of understanding musical abstractions.

2. "Sight-singing" and "Sight-reading" are frequently used to identify the same kind of music reading activity, the ability to accurately reproduce these music symbols at first sight through a vocal response. This interpretation is entirely logical if "sight-singing" is used. It would appear more logical to reserve the use of the second term for reference to situations where these symbols are accurately reproduced at first sight through performance on an instrument. In either case, it does not seem that the term "music reading" is sufficiently direct or specific.

3. Music reading may be considered as the process of reading and interpreting the various kinds of music symbols and converting these symbols into sound. This does not imply accurate reproduction at first sight but does suggest that the individual needs to acquire sufficient skill in understanding and interpreting the meaning of the symbols so as to ultimately reach a high level of correct response. The symbols become a guide for action, with the individual making frequent value judgments regarding the accuracy of his performance, and consequently changing that performance to agree with these judgments. Many music teachers have interpreted the music reading program as a program designed to develop sight-singing ability rather than the more flexible, utilitarian interpretation suggested here.

The present study accepts music reading as defined in Item 3, with the level of skill in music reading measured in terms of a vocal response to the stimuli. If an instrumental response is used, it is difficult to judge whether the subject really understands and interprets the musical symbols being read, or whether the response is mechanical and is being made in terms of fingerings. Vocal response was also considered to be the most practical for this study because all of the children had had extensive experience in singing as a part of the regular music program.

The Purpose of the Study

The general purpose of this investigation is to study the performances of selected fourth and sixth grade pupils as they engage in certain music reading activities. The primary purpose is to determine whether there are significant differences between normal children and children gifted musically in the ways in which they perceive and respond to the music symbols commonly associated with music reading. The investigation will also seek to determine whether there are significant differences between fourth and sixth grade pupils in each of these ability groupings. Within the limitations of the study, it may also be possible to compare two teaching procedures commonly employed in the development of music reading skills. The entire investigation is to be considered an exploratory study in this problem area, with the results being used to suggest research techniques and procedures which may be applied to other aspects of the problem.

The present study seeks to develop experimental learning situations which would provide significant information regarding the perception of music symbols by children of average musical ability and children gifted musically at the fourth and sixth grade levels. The process of reading music appears to depend upon three perceptual levels: (1) the auditory perception of musical sounds; (2) the visual perception of musical symbols; and (3) the integrative, internalized process through which the individual organizes previous auditory and visual perceptions of given stimuli in order to react to these same, or similar, stimuli as they appear in new learning situations. Each of these three perceptual levels will be explored experimentally, comparisons will be made in terms of the two groups of children at each of the two grade levels, and hypotheses will be tested.

The Procedures of the Study

The procedures followed in the study are briefly outlined below and will be discussed in detail in a later section of the report.

1. Analysis of song materials to identify common tonal configurations which may be used as test items.
2. Selection of subjects for the study and collection of certain personal data on each subject.
3. Construction of an initial test and administration of this test to subjects selected at random from the fourth, fifth and sixth grade pupils. This constitutes Phase I of the study and will provide one means of examining perception as the visual and aural identification and recognition of tonal configurations. It will also establish, in part, the basis for selecting learning tasks appropriate to Phase II of the study and provide an opportunity to study the techniques appropriate for the collection and analysis of the data.
4. Construction of the test materials for Phase II of the study and administration of the test to average and musically gifted children in both fourth and sixth grades. Phase II is primarily concerned with perception as an integrative, internalized process through which the individual relates his auditory and visual perceptions of tonal configurations to his perception of these same configurations as they

appear in a different setting.

5. Final analysis of the data, with small group comparisons between grades and between groups being made.

The testing carried on in the study is all individual, with the data gathered by tape-recording the pupil responses in the testing situations and later subjecting these recordings to detailed analysis. Tape recording the pupil responses in the testing situation eliminated the need for immediate scoring and tabulating and insured accuracy of the data since questionable responses could readily be repeated for verification.

The hypotheses of the study are as follows:

1. Normal children perceive and act upon music symbols in essentially the same way as do children gifted musically.

2. Certain tonal configurations which are readily perceived in isolation by both groups of children tend to retain their identity when presented in the more usual context of a song.

3. There are no significant differences between the ability to perceive a tonal configuration aurally and the visual perception of these same configurations.

4. Age, previous musical experience and musical ability are significant factors in developing skill in perceiving and reacting to music symbols.

SUMMARY OF RELATED INFORMATION

The writer has made an extensive analysis of the professional literature relating to the basic problem of music reading. This analysis will be summarized briefly at this time although it should be understood that it will include only those reports considered relevant to the present study. The problem of developing skill in reading music has been a fertile topic for discussion in the professional literature for a period of many years. Music educators have offered many suggestions and recommendations with reference to appropriate procedures for the development of this highly specialized skill. The objectives of the music program in the schools, as stated by authorities in the field, have always mentioned the need for acquiring a certain level of music reading skill. There exists, however, a considerable amount of uncertainty regarding exactly why children need to learn how to read music, as well as a general tendency to avoid making specific recommendations regarding affective and adequate procedures which might be followed. One valid reason for the vagueness with which these two points are discussed may well be due to the lack of adequate research information.

The literature concerned with the topic of music reading may be grouped into three broad categories:

1. The historical background of musical notation and of syllabic devices for interpreting musical notation.
2. Professional literature which includes courses of study, specialized music education methods books, and the many articles in professional journals which report the opinions, experiences, and recommendations based upon personal experience.

3. Research literature.

This report will not summarize the literature available in the first category because, despite the historical significance of the problem, the writer does not consider it relevant to the report. The literature available in the second category will not be summarized completely for the same reason. An attempt will be made, however, to present an adequate summary of pertinent research literature.

Purposes of the Music Program

One very succinct statement of the purposes of the music program in the schools may be found in an early book by a recognized leader in the field, Mursell:

> "The best way to insure the place of music in the schools is to make it humanly significant. Let us try to help the children to learn to sing in school, songs they will love to sing outside. Let us give them authentic and moving experiences in the way of listening to great music. Let us point our program towards the actual making of music by participation in musical organizations. Let us bring to them the joys of bodily freedom so essential to fine musical perceptions, through creative work in rhythm. Let us encourage them to compose music for themselves and in this manner find new avenues of self-expression. Let us carry on all that we do in an atmosphere of joy in musical achievement." (9:21)

An equally prominent authority in music education, Pitts, offers this suggestion regarding the purposes of music education:

> "If the aims and purposes of music education are identical, ultimately, with the aims of education in general, music is but another, though powerful means of making a difference in the way children conduct themselves and their lives; what they are inside of themselves and what musical expression can do to enlarge the personalities and enrich the social living of __all__, not just a __few__ children." (14:44)

In a recent publication of the National Society for the Study of Education, Burmeister calls attention to the unique contributions which music makes to the general education of the individual in the areas of (a) aesthetic growth; (b) productive use of leisure time; and (c) emotional development. He cautions against supporting its place in the curriculum by treating music as an instrument for the attainment of goals which are not unique to music, such as health, social growth, citizenship training, etc. While music is certainly not the only activity related to emotional development, it possesses these attributes:

> "1. Music is the most subtle, pervasive, and insistent of all the arts.
> 2. It requires not intellectualization to work its effects.
> 3. Its effects cannot be denied by the auditor. It is impossible to direct or divert the psychophysical effects of the sheer potency of tone by an act of will.
> 4. Special abilities are not necessary. All can share in a response to music." (2:220)

A final summary of the purposes of the music education program is to be found in the official statement of the Music Educators National Conference.

> "The function of music in general education should be not only to provide opportunity for the gifted child to experience and perform music at the level of his capabilities, but should also provide for the enrichment and development of the personalities of all students through their heritage as citizens of America and the world. Music is needed to contribute to the rounding out of the whole child as a process of his natural development and should be a part of his daily experience.
>
> There is some phase of music in which every child can succeed. Equipment and opportunity for him to feel success must be provided. The child needs to be given confidence in every possible music activity." (7:15-16)

These general statements all tend to support the contention that music should, and can, serve a unique function in the school of today. With regard to the purposes of music reading in such a program, the writer has selected the following statements as representative of the current thinking of music educators.

Purposes of Music Reading

Brooks and Brown, as well as others, feel that "mastery of the musical score" is a broader, more appropriate term than learning to read music. "The score is a system of symbols used to represent musical ideas, and the mastery of the score for the child should include both the reading of the score and its use in expressing his musical ideas." (3:176) They go on to say that "the best reasons for a mastery of the musical score in the elementary school are (a) to open to the child a vast storehouse of music literature which would otherwise be closed to him and (b) to put him in possession of a means of musical expression." (3:176) For these reasons the authors feel that all children can and should learn to read music, exceptions being children who are organically defective. The authors go on to present a lucid discussion of the problem which emphasizes that the school must provide music activities which will stress both the need for ability to read music as well as opportunities to gain the kinds of experiences from which such skills will emerge. Myers agrees with this point-of-view when she states: "Children need power of their own in interpreting the symbols of music. Independent ability in translating symbols is the goal of music-reading activities." (10:178) She feels that musical experiences of the school should be designed so as to bring about a keen desire on the part of the child, to investigate further the skill of reading music. Throughout her presentation is the constant reiteration that music reading is taught by using combinations of symbols, rather than by individual symbols.

Mursell gives similar reasons for developing music reading skill:

> "A few musical concepts can be conveyed in nontechnical language. But in most cases a special system of symbolism is far more convenient, and often it is indispensible......The standard notation, in spite of its many anomalies, is our best and most adequate means of symbolizing musical concepts......It is altogether necessary that these symbols be learned. Otherwise, musical development is bound to remain at a low level, and musical apprehension to be vague, crude, relatively incoherent...... They (the symbols) must be taught always in terms of their musical meanings and in application to musical situations and experiences, never merely in terms of verbal definitions and arithmetical designations." (8:153)

> "The musical score is a visual medium giving precision to the auditory impressions which are the essential stuff of music......It can reinforce and render more definite the mental processes, not only of performance but also of listening and creation......Musical symbolism is a valuable aid to musical thinking, feeling, and imagining, although its mastery will not, in and of itself, create in anyone the ability for such thinking, feeling, and imagining.......A mastery of the score liberates the child for an advance to more complex and exacting musical undertakings......The notation should always be regarded as an instrumentality through which one may grow in refined and precise musical power." (9:182-183)

In summary, the authorities appear to agree that music serves a unique function within the overall school program and that mastery of the musical score, or skill in music reading, is essential for the child's musical growth. The music program needs to present interesting and varied activities which will create within the child a desire to want to learn to read music, activities which will create a need for such a skill. These very excellent generalizations are accepted by most music educators. However, the problem of developing skill in music reading still exists because of the uncertainty which seems to accompany all efforts to implement the recommendations. This uncertainty exists because the recommendations are based largely upon practical experience and lack the support which comes from significant research activity.

Research Literature

Analysis of the research in the specific area of music reading reveals the following pertinent information. One group of studies has approached the problem through a detailed analysis of the eye movements of persons engaged in reading music. Jacobsen (6) following the procedures of Gray and Buswell in the field of language reading, carried forward a detailed photographic analysis of eye movements as a means of identifying the various stages in the development of music reading ability. The section of the report which is pertinent to the present study deals with the reading of vocal music by three college students and three sixth-grade pupils. A detailed analysis of the performance of each subject is presented and the results are briefly summarized here. The less mature reader required pauses almost twice as great as the mature reader but the number of pauses did not differ significantly. The span of recognition, which included both words and notes, seldom exceeded two notes for the mature reader and was smaller than the span of the earlier instrumental reading. The smaller intervals were read correctly more often than longer intervals; use of tonic sol-fa syllables resulted in longer pauses than when words were used; and sol-fa syllables were found to be of no benefit to the subjects. The author offers several suggested applications of the results of this study to classroom instruction which seem to indicate that new and simplified reading materials are needed to develop the immature reader.

Ortmann (13) summarized one of his investigations by pointing out that the span of vision depends a great deal upon the placement and pattern of the notes. He found the field of clear vision to be a circle varying between 1/4-and 1/3-inch in diameter with no difficulty encountered in reacting to notes within this area. The chief determinants of difficulty in note reading, not involving actual eye movements, are: (a) the number of notes in the field of clear vision, (b) the area covered by the distribution of the notes, (c) the number of linear dimensions involved (i.e., combinations of vertical, horizontal or oblique), (d) the complexity or symmetry of the note pattern, and (3) the meaning of the note group as a melodic and/or harmonic pattern. He points out that difficulties in note reading, for the normal pupil, result from an inability to group a distribution of notes into larger perceptual units. Familiarity with musical notation is quite obviously a significant factor in accuracy of reading.

A third study of eye movements in music reading was carried on by Weaver (17) using trained musicians, with the piano as the performance medium. His experimental reading materials; harmonic, melodic, and single melody with supporting chords. The rhythmic element was retained in all types as a means of accurately measuring the number and duration of reading pauses. The results indicate that note reading relates more closely to reading algebraic formulae than to language reading, assuming that one musical note is equivalent to one word. The average number of notes played per reading pause varied between one and two, with spacing of notes and harmonic grouping being important factors. The duration of reading pauses depended upon the type of musical material being read, with those for harmonic and melodic material reliably shorter than those required in reading an accompanied melody. A higher correlation exists between reading time and pause duration than between reading time and the number of pauses. Weaver also found that performance in his study provided little basis for distinguishing types of readers because of the individual differences in performing the different selections. It is significant that he identified the average perceptual span as varying between three and five notes. A companion study by Weaver and Van Nuys was particularly concerned with the relationships between the rhythmic and melodic elements of music as they affected occular and manual behavior. The immediate memory span was used as a criterion of difficulty in determining the eye-hand span, with immediate memory span defined as the amount of material which could be visually explored, perceptually organized and retained until played. The results showed that memory span decreases as the complexity of note relations increases for both rhythm and melody. The memory span for a combination of a simple rhythm and melody is only slightly smaller than the span for each alone. In more complex combinations, the melodic element seems to prove more difficult than the rhythmic. Melodic factors limit memory span whenever the melody is not extremely simple; rhythmic factors limit the rate of reading when complex rhythms are used. If music reading skill is to be developed so more rapid responses are made, individuals need to improve in the ability to apprehend pitch patterns as stable melodic segments and to grasp rhythms as rhythmic figures.

Bean (1) sought to determine the complexity of the musical pattern which could be perceived at one fixation of the eyes by individuals with varying amounts of training. He also studied the effects of tachistoscopic training on the span of perception. The investigation showed that the ability to grasp a musical pattern of three or more notes at a glance was essential to rapid, accurate reading. The perceptual unit is determined by the duration of notes, the distribution of notes on the staff, the closeness of note groupings, and the irregularity or unusual characteristics of patterns. These factors all operate to determine the ease with which material is grasped by the reader. Disjunct melodic patterns, rhythmic patterns which could not readily be subdivided into logical sub-groups, unfamiliar musical material, and uneven distribution of chord tones appear more difficult to perceive than other types of material.

Christ (4) employed tachistoscopic training as a means of improving the ability to perceive and reproduce rhythmic units. The results showed that such drill yielded significant gains but no attempt was made to determine the permanence of such drill or to relate such drill to rhythmic reading of typical music. Stokes (16) investigated the effect of tachistoscopic training on reading melodic material. The results showed, however, that the gains on a post-test could not be considered due to the influence of the specific training program alone. An interesting instrument for testing the sight-reading performance of instrumental performers was constructed by Stokes and contains several valuable suggestions. The testing results show, however, that tachistoscopic training had no effect upon general achievement in reading music and that there was little evidence of carry-over.

Smiley (15) investigated the music reading "vocabulary" in basic music texts at the fourth grade level and found the total number of tonal configurations to be far too many for the child to respond to readily. The 351 configurations could, on the basis of frequency of use, be reduced to approximately 20, each containing from two to five notes. Further analysis led Smiley to conclude that it might be more valuable for the child to learn a basic list of two-note configurations of seconds and thirds. The procedures suggested by Smiley for identifying tonal configurations were a valuable guide for the analysis made by the writer.

This reported research in music reading has been primarily concerned with certain aspects of visual perception or of physical behavior during the process of reading music. Such research has not yet provided an adequate basis for prescribing instructional procedures to develop musical concepts which may be readily utilized in a total musical setting. Apart from a few reports of actual classroom experimentation in the problem area, there is little information regarding the ways in which children of elementary school age react to musical notation because the majority of the studies have used older subjects. The survey of literature again reemphasizes the need for continued, systematic investigation of this, and other basic problems in music education.

ANALYSIS OF SONG MATERIALS

The first step taken in the present study was the analysis of basic song-texts to identify common tonal configurations. The purpose of this analysis was to select possible items for use in the testing situations. Since the subjects for the study were to be selected from the fourth, fifth and sixth grades, it was necessary to use items with which they would be reasonably familiar and still have items which would be discriminating. It was considered logical, therefore, to select items on the basis of frequency of appearance within these song materials. In order to have test items with which most children would have had considerable aural familiarity, it was further decided to exclude tonal configurations from songs in a minor mode, or those which were found in chromatic or modulatory passages.

The procedures followed in identifying the tonal configurations were suggested, in part, by Smiley. (15) Other guides in establishing procedures to be followed were obtained from the teachers' manuals to the basic song-texts which listed common configurations in the particular series, and from professional books whose authors advocated the "pattern" approach to music reading. Eighteen of the basic song-texts used in grades three through six were chosen for analysis. The texts are listed in Appendix A. A final total of 326 songs, which had been selected at random from these books, were analyzed to identify common tonal configurations. The following rules were set up to serve as a guide in defining a tonal configuration:

1. Thirty page numbers, chosen from a table of random numbers, were set up for each of the eighteen song-texts. Although only eighteen songs in each book were to be analyzed, the additional songs were

necessary as alternates because of the decision to eliminate songs in a minor mode or songs containing frequent chromatic or modulatory passages.

2. No distinction was made between specific reading songs, rote songs, or rote-note songs and all types were included in the analysis.
3. Single intervals were generally not considered as tonal configurations since they do not imply any melodic flow or movement.
4. The interval between musical phrases was not usually included.
5. The interval within a phrase which appeared between the end of one distinct configuration and the beginning of another configuration was not included unless the text showed that such an interval occurred in the middle of a word.
6. The interval between readily identifiable sequences was not included, each sequence being treated as a separate configuration.
7. Repetitions of single notes were considered as single notes; do-do-re-re-do-do- was analyzed as do-re-do. The decision to exclude repeated tones from configurations was made in order to simplify the analysis even though such repetitions occur frequently in songs.
8. In part songs, each part was analyzed separately in those sections of the song where the soprano and alto parts were different. In sections where the soprano and alto parts moved in unison, only a single analysis was made.
9. In songs containing repetitions, only the new ending was analyzed, not the entire repetition.
10. Frequently the rhythmic structure within the measure or phrase could be used to assist in the identification of a complete tonal configuration.

11. The text of the song itself, particularily with reference to words of more than one syllable, could frequently be depended upon to be of assistance in accurately identifying the tonal configurations.

12. Within these guides, the decision regarding the length of the configurations had to be made by the investigator in terms of what the child could be expected to grasp. As a result, tonal configurations in excess of five or six tones were rarely identified as such.

The writer evolved a system of entering and tabulating the configurations which proved to be quite satisfactory. All items were arbitrarily transposed to the key of E flat major and entered on data sheets in musical notation. Individual data sheets were used for tonal configurations in each of the following categories:

1. Tonal configurations based on the tonic chord.
2. Tonal configurations based on the subdominant chord.
3. Tonal configurations based on the dominant or dominant seventh chord.
4. Tonal configurations beginning on each of the several tones of the major scale. These were sub-divided into ascending configurations, descending configurations, neighboring-tone configurations, and conjunct or disjunct configurations containing intervals of a third or larger.

To simplify the report, however, tonal configurations will be identified by their tonic sol-fa syllables rather than in acutal musical notation.

The complete analysis of all tonal configurations identified in the songs will not be presented in this report since much of it is not pertinent to the study. A partial summary of the data will be sufficient to establish the bases for the selection of test items.

TABLE 1

Table Showing the Frequency Count of 558 Tonal Configurations

Frequency Range	Number of Configurations	Per cent of Total Count	Combined Frequency	Per cent of Total Frequency
44-213	22	3.94	1939	42.38
20-43	25	4.48	646	14.12
10-19	51	9.14	663	14.49
1- 9	460	82.44	1327	29.01
Totals	558	100.00	4575	100.00

Table 1 summarizes the numerical results of the analysis. The total number of different tonal configurations identified in the 326 songs was 558. The total frequency of these 558 configurations was 4575, with a frequency range of 1 to 213. It is significant that 47 tonal configurations, or only 8.42 per cent of the total number identified, with a frequency range of 20 to 213, had a combined frequency count of 2585, or 56.50 per cent of the total frequency count. Even more significant is the fact that 22 configurations, or not quite four per cent of the total, with a frequency range of 44 to 213, had a combined frequency of 1939, or 42.38 per cent of the total frequency count. Tonal configurations with a frequency range of 1 to 9, accounted for 460, or 82.44 per cent of the total number of configurations, but had a combined frequency count of only 1327, or 29 per cent of the total frequency count. The data points up the fact that there exists a relatively small number of tonal configurations which appear repeatedly in the songs children sing.

Had songs in a minor mode been analyzed, as well as songs containing chromatic and modulatory passages, the total number of tonal configurations would have been considerably in excess of 558. The situation is still further

complicated by the fact that each one of these 558 configurations can be written in at least nine different major keys commonly used in these song-texts. Since each configuration has, therefore, nine different notational systems within which it could appear, the total possible number of 5022 configurations appears to present a very formidable problem to the music educator faced with the task of identifying effective methods for developing skill in music reading! Fortunately, the general outline or shape of a configuration remains constant from one tonality to another despite differences in actual notation. However, this problem of perceiving and identifying configurations when they are presented in several different tonalities needs careful investigation since it is frequently dismissed as being relatively unimportant.

Since the selection of test items was to be based, in part, upon the frequency of appearance in songs, the twenty-two most frequently used configurations are listed in Table 2, page 24. For purposes of comparison, the twenty-two most frequently used configurations identified in the analysis made by Smiley are also presented. These twenty-two configurations represented 3.94 per cent of the total number of configurations identified by the writer and 6.1 per cent of the total number of configurations identified by Smiley. It is significant to note that although Smiley confined his analysis to fourth grade songs identified as reading songs, his findings closely agree with those of the present investigator. Nineteen configurations are common to both lists, while each list contains only three configurations not found in the companion list. These configurations are identified by an asterisk. It would be reasonable to expect analyses of this type, carried on by two investigators operating independently but following similar criteria of identification, would show comparable results. However, since different songs were selected for each of the analyses and since accurate identification of tonal configurations probably should be made by some other method, one could not reasonably anticipate that a straight-line relationship would exist between the two lists if items are considered individually in terms of their rank. In fact, a rank-order correlation, using the formula $r = \frac{6\Sigma D^2}{N(N^2-1)}$, was calculated resulting in a relatively insignificant "r" of .202.

In order to insure an adequate number of tonal configurations from which test items could be constructed, an additional 25 configurations were listed. These configurations, with a frequency range of 20 to 42, accounted for 646, or 14.12 per cent of the total frequency count and are listed in Table 3, page 25. It is significant to note that there again exists considerable agreement between the analysis made by Smiley and that made by the writer. Sixteen, or 64% of these 25 configurations appear on both lists. The 47 tonal configurations, with a frequency range of 20 to 213, which have been identified by the writer, included 35 configurations having a similar frequency range as identified by Smiley. This represents a concurrence of identification for 75 per cent of the highest ranking items. No attempt was made to compare results of this analysis with the tonal configurations identified in professional books and teacher's guides to the basic song-texts since it was not considered significant to this particular study.

These 47 configurations were considered a sufficient number from which to select test items and any further analysis of the data was unnecessary at this time. The analysis was of considerable value in emphasizing a basic problem which the writer feels needs clarification if music educators are to continue to support the "pattern" approach to music reading. It involves establishing objective, rather than subjective criteria for accurately identifying tonal configurations. Such objective criteria might well be established

Table 2

Table Listing the 22 Most Frequently Used Tonal Configurations as Identified in the Present Study and by Smiley.

Present Study			Smiley's Study		
Configuration	Frequency	% total Freq.	Configuration	Frequency	% total Frequency
so do	213	.047	So do	191	.041
mi re do	167	.037	so la so	135	.029
so la so	136	.030	do so	102	.022
mi fa so	115	.025	mi re do	99	.021
do re mi	110	.024	so mi	91	.020
do ti la so	106	.023	mi so	88	.019
do mi so	93	.020	do ti la so	80	.017
so mi do	93	.020	do ti do	73	.016
so fa mi	92	.020	do re do*	72	.016
so mi	90	.020	so mi do	70	.015
mi so	76	.017	do mi so	69	.015
so la ti do	76	.017	fa me re do	69	.015
do la so	69	.015	do re mi	66	.014
do so	68	.015	mi do	65	.014
do ti do	65	.014	so fa mi	64	.014
so fa mi re do	60	.013	do la so	60	.013
so mi so*	58	.013	so la ti do	59	.013
mi do	55	.012	mi fa so	58	.013
do mi	55	.012	so fa mi re do	58	.013
fa mi re do	54	.012	so la*	53	.011
so do mi*	44	.010	do la*	46	.010
re mi fa*	44	.010	do mi	43	.009
Total	1939	.424	Total	1711	.369

if the procedures suggested in the research concerned with the photographing of eye-movements were to be applied to the problem. Certainly such an investigation would also need to consider the ways in which the perception of any given tonal configuration might be significantly influenced by the factors of rhythmic structure, the accompanying text of a song line, and the overall movement and direction of the melodic line.

TABLE 3

Table Showing the Frequency Arrangment of 25 Tonal Configurations
Having a Frequency Count of 20 to 39

Configuration	Frequency	Configuration	Frequency
* do re mi fa so	39	* so mi	25
do re do	36	* so do	24
* do re mi do	31	so fa re do	23
* mi fa mi	30	* fa re	23
* fa mi re	30	* la ti do	23
do do	29	do la do	22
* do ti la	28	mi so do	22
* re mi re do	27	so la so fa mi	22
* do mi do	27	* la so fa mi	21
* mi re mi	27	so do so	20
* do re mi re do	26	re fa mi re do	20
do so do	26	* so fa mi re	20
mi so mi	25	Total	616

*These items were also identified by Smiley as having a frequency count within this range of 20 to 39

SELECTION OF THE SAMPLE

Preliminary Statement

The original design of the study called for a population representing the fourth, fifth and sixth grades of the Madison, Wisconsin, public elementary schools. The sample was to have been selected at random from this population. Arrangements were made with appropriate members of the administrative staff of the Madison school system to carry forward an extensive program of individual testing. It was mutually agreed that all such testing would be done during the subject's regular music class thus eliminating the problems which often arise when pupils are taken out of regular academic class periods. It soon became apparent that a sample from all of the elementary schools was impossible to derive because of conflicting class schedules, lack of available testing space in certain of the schools, the amount of time required to select such a sample, and the lack of sufficient qualified research assistance. It was therefore decided to utilize only three out of the total twenty-two schools, these three being selected as representative of the total school population.

Selection of the Population

There are 22 public elementary schools in the city of Madison and three of these were selected as being fairly representative of the overall school system. The criteria for selection were school size, geographic location and socio-economic background of the children. Recommendations regarding the selection were secured from the supervisor of music, several elementary school principals, and representatives of the central administrative staff.

With regard to the first criterion, size, it was decided to select one school from each category of small, medium and large. Table 4 shows the enrollment figures as of December 20, 1958. The total enrollment for all 22 schools was 12,055.

The mean enrollment for these 22 schools is 548 and the median is 410. Because of its location, the city of Madison may be divided into three geographic areas; central, east and west. There are seven schools in the central area, accounting for 3504, or 29.1 per cent of the total elementary school population. Six schools in the east area account for 4417, or 36.6 per cent of the total; while nine schools in the west area account for 4134, or 34.3 per cent of the total. One school was selected from each area. The criterion of socio-economic background was more difficult to establish because the geographic areas do not clearly define such levels. It was felt, however, that adequate differentiation would result from selecting a school representing each of the following: professional group; semi-professional group; and skilled and semi-skilled group. Application of these three criteria made it possible to select the schools from which the sample would be derived.

School A is a large school, enrollment of 841, located in the east area of the city where pupils are drawn largely from the semi-professional group. School B is a small school, with an enrollment of 478, located in the central area of the city where pupils came from the skilled and semi-skilled group. School C is of medium size, enrolling 660 pupils, and located in the west section of the city in an area largely composed of the professional group.

Table 4

Table Showing the Relative Sizes of the 22 Elementary Schools in Madison, Wisconsin

Enrollment	Number of Schools
200--299	1
300--399	4
400--499	7
500--599	2
600--699	4
700--799	1
800--899	2
Over 900	1
	22

Table 5

Table Showing the Distribution of the Subjects Participating in Phase I.

SCHOOL	4th Grade		5th Grade		6th Grade		
	Boys	Girls	Boys	Girls	Boys	Girls	Totals
School A	5	5	5	5	5	6	31
School B	4	5	4	5	5	5	28
School C	5	6	5	4	5	5	30
Total	14	16	14	14	15	16	89

Selection of Sample for Phase I

The study is divided into two distinct parts, each utilizing a different sample. Phase I utilized a "normal" population of pupils selected at random from the fourth, fifth and sixth grades of each of three schools and undifferentiated in terms of musical ability. Class lists were secured and an equal number of pupils from each grade in each school were selected at random, with equal numbers of boys and girls. Although more than 100 children were tested in Phase I, usable data was secured for 89 pupils. Table 5 shows this distribution by school and by grade.

The data for several subjects was judged unusable because they were unable to make the type of response called for by the test, a singing response to aural and visual stimuli. It was considered unwise to administer the entire test to children for whom there could be little success on any item.

Selection of Sample for Phase II

1. Procedures Used to Identify Musically Gifted Children

Phase II of the study utilized two reasonably well-defined segments of this "normal" population of the three selected schools, pupils identified as those gifted musically and those possessing average musical ability. Subjects were selected from grades four and six for each of these two groups. The initial step taken in the identification of these two groups was to establish criteria which could be followed. There is considerable discussion in professional literature regarding the gifted child but there exist few criteria for identification in the field of music. Usually such judgments are made on the basis of one or more of the following: (a) excellence in musical performance; (b) demonstrated knowledge of musical information and concepts as measured by achievement tests; (c) results of music aptitude or talent tests; (d) grades in music subjects; and (e) rating scales constructed for the purpose. The schools participating in this study do not follow a standard testing program to establish levels of musical achievement although a short paper-and-pencil test for knowledge of musical notation is given to all sixth-grade pupils. There was, therefore, no readily available data which could be used for identifying possible gifted children. An investigation of available professional literature had failed to yield many concrete suggestions regarding procedures which could be used to identify musically gifted children other than the criteria mentioned earlier. Each school had an experienced special music teacher and it was felt that this person would be capable of identifying gifted pupils if certain criteria were made available. Consequently, a rating sheet was designed by the writer and included twelve items considered by music educators as important factors in determining musical accomplishment of elementary school pupils. The rating scale was submitted to a group of music teachers for comments and suggestions and the final form was drafted. A copy of this rating scale may be found in Appendix B. The music specialist in each school was then asked to evaluate all fourth and sixth grade pupils and the rating sheets were then scored and tabulated. Obtaining evaluations for all pupils made it possible to identify for each grade in each school, three groups of pupils: (1) those receiving the highest scores; (2) those receiving the lowest scores; and (3) those in the middle of the distribution. This procedure led to a more accurate identification of the "average" group since it could be assumed that they were to be found between the upper and lower groups. Table 6 presents a summary of these teacher ratings. For each grade in each school the upper ten percent of the total distribution was

Table 6

Table Showing the Mean Average Scores and Standard Deviations of the Teacher Ratings.

School	4th Grade			6th Grade		
	Number	Mean Average Score	S.D.	Number	Mean Average Score	S.D.
A	96	3.58	.955	112	3.72	.951
B	70	3.48	.807	50	3.45	.825
C	79	4.29	2.83	68	3.84	.974

(Highest possible mean score is 5.0)

Table 7

Table Summarizing the Scores Earned on the Kwalwasser Test by 438 4th and 6th Grade Pupils in Three Schools

School	4th Grade			6th Grade		
	Number	Score Range	Mean	Number	Score Range	Mean
A	94	19-36	26.88	106	13-37	28.73
B	70	19-35	27.50	50	21-40	28.58
C	54	20-36	28.37	62	17-37	29.66
Total	218	19-36	27.45	220	13-40	28.96

(Highest possible score is 40.)

tentatively selected as being musically gifted. The mean rating score and standard deviation was calculated for each grade in each school and the "average" group was tentatively defined as being within the limits of one-half a standard deviation on either side of the mean. The ratings for each school were kept separate since it obviously could not be assumed that all three teachers would interpret the criteria in the same way.

In addition to the teacher ratings, a second criterion measure was used to identify the musically gifted. The Kwalwasser Music Talent Test, Form B, was administered to all fourth and sixth grade pupils in the three schools. The decision to administer this particular test was based on the practical necessity of having a standardized talent test which could be given within a single twenty-five minute period, the length of the music period in the schools. The writer would have preferred a test for which higher validity and reliability had been established, but the time limitations precluded the use of a longer test. The test was administered to 438 pupils and the data is summarized in Table 7. The mean score for each sixth grade is slightly higher than the mean score for each fourth grade, but there appears to be little significant difference. It was not considered necessary to test for the significance of the difference at this time. For the combined grades in the three schools, with a total N of 438, the range of scores was 13 to 40 and the mean was 28.21.

A comparison of the data obtained from the teacher ratings and the Kwalwasser test shows that a relatively low correlation exists between the two criterion measures. Combining all pupils in both grades in each school and calculating the coefficient of correlation for each school showed the following: (1) School A, $r = .34$; (2) School B, $r = .39$; (3) School C, $r = .34$. In the absence of other information, one may conclude either that teacher ratings provide a more valid way of determining musical giftedness than do talent tests, or that the reverse is a more appropriate conclusion. The writer decided to consider this aspect of the problem in more detail after the data had been collected for Phases I and II. Judgments regarding those pupils who were to be classified as musically gifted were made on the basis of arbitrarily assigning equal weight to both the teacher ratings and the Kwalwasser test scores and then selecting pupils who ranked highest on both measures.

2. Methods Used to Secure a Representative Sample

In order to obtain a more representative sample for Phase II, it was decided to select the sample from each school on the basis of school and class size. Table 8 gives a summary of the enrollments in each grade in each school. An arbitrary total of 150 subjects was established as the maximum number which could be tested within the time limits of the study. The data shows a total of 475 pupils of which 245, or 52 per cent were fourth grade pupils and 230, or 48 per cent were sixth grade pupils. School A accounted for 208, or 44 per cent of the total; School B for 120, or 25 per cent; and School C for 147, or 31 per cent. These proportions were maintained in the sample, as well as the proportionate representation for each grade in each school. Table 8 also shows that for School A, 96, or 46 per cent of the total number of pupils were in the fourth grade while 112, or 54 per cent were in the sixth grade. Therefore, the total sample from School A was 66, or 44 per cent of the total sample of 150, with 30, or 46 per cent of these 66 to be fourth grade pupils and 36, or 54 per cent of the 66 to be sixth grade pupils. The same process was used to derive the other samples. The number of gifted pupils identified by the two criterion measures were then subtracted from the total number allocated to each grade in each school and the remainder constituted the number of average pupils which were to be selected at random from the distribution within the limits of one-half a standard deviation on either side of the mean. Additional subjects were selected in order to insure the availability of alternates in the event a selected subject was unable to participate in the testing program. Table 9 gives the number of average and gifted pupils initially selected for the Phase II of the study.

Table 8

Table Showing Pupil Enrollments in 4th and 6th Grade for each of the Three Schools.

	4th Grade		6th Grade		Total	
SCHOOL	N	Percent of Total for the 2 Grades	N	Percent of Total for the 2 Grades	N	Percent of Total for the 3 Schools.
A	96	46	112	54	208	44
B	70	58	50	42	120	25
C	79	53	68	47	147	31
Total	245	52	230	48	475	100

Table 9

Table Showing the Representative Sample of Pupils Initially Selected for the Testing Program of Phase II.

SCHOOL	Total Subjects	Percent Of Sample	4th Grade		6th Grade	
			Average	Gifted	Average	Gifted
A	66	44	20	10	21	15
B	38	25	15	7	8	8
C	47	31	14	10	13	10
Total	151	100	49	27	42	33

PROCEDURES OF THE STUDY

Phase I

1. Purposes

The purposes of Phase I of the study were: (1) to examine perception as the visual and aural identification and recognition of the similarities and differences between tonal configurations by children in fourth, fifth and sixth grades; (2) to study the techniques appropriate for the collection and analysis of the data; and (3) establish, in part, the basis for selecting learning tasks for Phase II.

2. General Plan

The testing procedures planned for Phase I were as follows:

1. The subject is given a visual presentation of the first tonal configuration, accompanied by some type of pitch indication to help him locate the starting tone.
2. The subject is then to sing an immediate response to this visual presentation by reading the notes of the configuration.
3. This reading trial is followed by a second visual presentation of the configuration in combination with an aural presentation of the same configuration. This permits the subject to associate the sound of a configuration with its visual representation.
4. Immediately after this aural presentation the subject sings what he has heard, continuing to look at the notation of the configuration.
5. The subject then proceeds to the next configuration, following steps 1, 2, 3 and 4. The testing continues in this manner until all the configurations have been presented.

6. The total list of configurations is presented a second and third time, with the configurations randomized for each successive presentation.

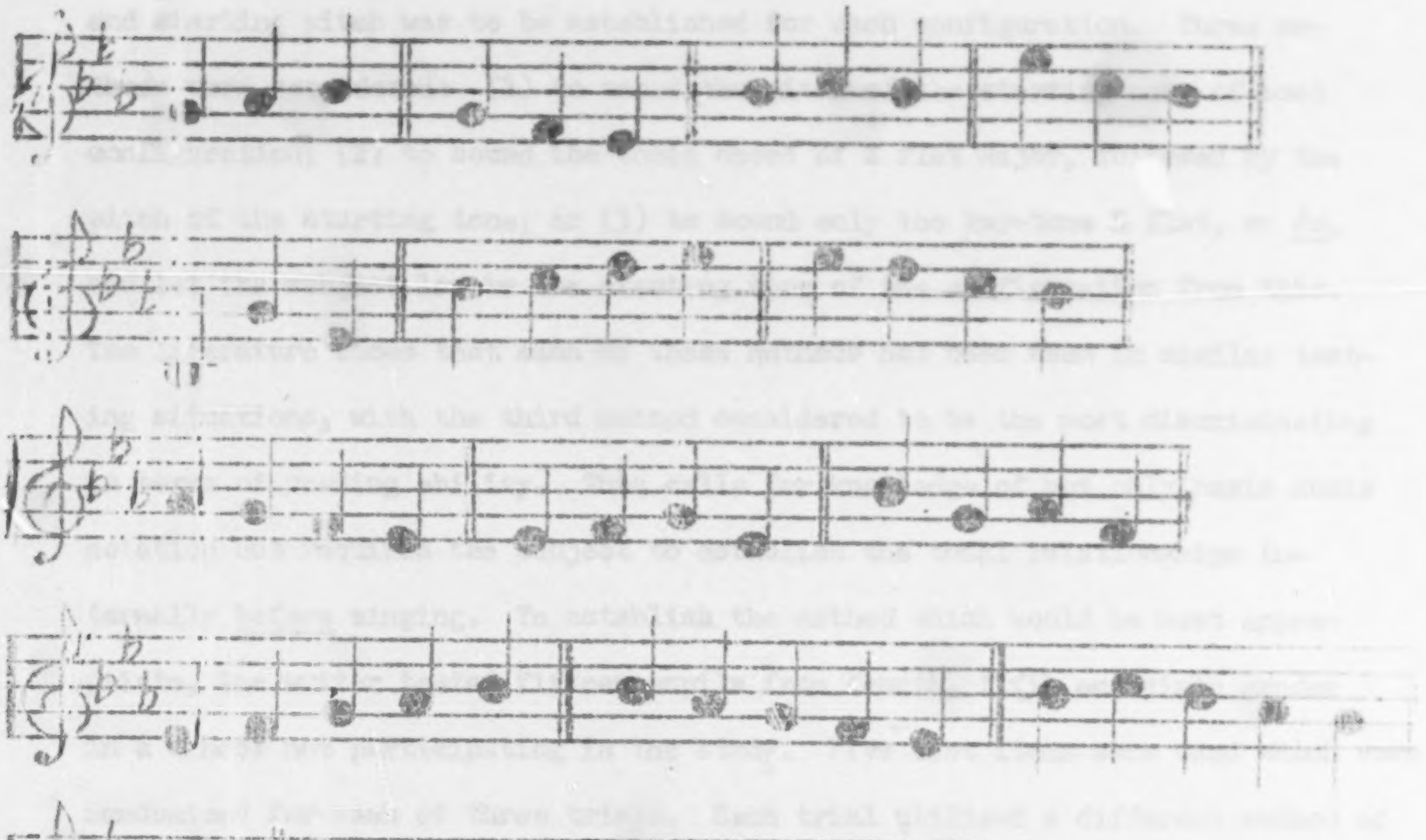
Since the testing is all individual and it is planned to tape-record the responses, the data should be of considerable value in attempting to determine the nature of the music reading process.

3. Construction of the Test

The decision to use a 20-item test was based, in part, on the assumption that several trials of the same item would provide data of greater significance for the present study than would single trials of many items. It was also felt that using only 20 items would keep the test relatively short and still present a variety of test items.

Twenty test items of varying length and complexity were constructed. These test items were selected from the tonal configurations which the analysis of song materials discussed earlier in Section III had identified as having the highest frequency. Equal numbers of three-, four-, five-, and six-note items were used and these represented ascending and descending scale and chord patterns, either presented separately or in combinations. The construction of these test items utilized some of the suggestions of Ortmann, who was interested in defining and measuring tonal memory in music. (11, 12) The items were written as quarter notes with no measure bars or time signature indicated. All the items were written in the same key to reduce the number of variables in the testing situation. The key of E flat major was considered the most appropriate because it would be possible to keep most of the items within a reasonable singing range for the children. The 20 test items are presented in Figure 1.

Figure 1
The 20 Test Items Initially Selected for Phase I.



4. Exploratory Studies

Several problems had to be solved before beginning the regular testing in Phase I. One of these was concerned with the manner in which the tonality and starting pitch was to be established for each configuration. Three methods were considered: (1) to sound the pitch of the starting note of each configuration; (2) to sound the tonic chord of E flat major, followed by the pitch of the starting tone; or (3) to sound only the key-tone E flat, or do, and let the subject locate the starting tone of the configuration from this. The literature shows that each of these methods has been used in similar testing situations, with the third method considered to be the most discriminating in terms of reading ability. This calls for knowledge of not only basic music notation but requires the subject to establish the tonal relationships internally before singing. To establish the method which would be most appropriate, the writer tested fifteen pupils from fourth, fifth and sixth grades in a school not participating in the study. Five test items were used which were randomized for each of three trials. Each trial utilized a different method of establishing the starting pitch of the configuration. Comments by the children, as well as test results, showed the third method to be too difficult, with a definite preference shown for the first method which gave the pitch of the starting tone of each configuration. This was then selected as appropriate for Phase I. The writer is aware that disagreement on this point could be justified but points out that the study was not primarily concerned with identifying those methods of giving pitch which could serve as a criterion of music reading ability.

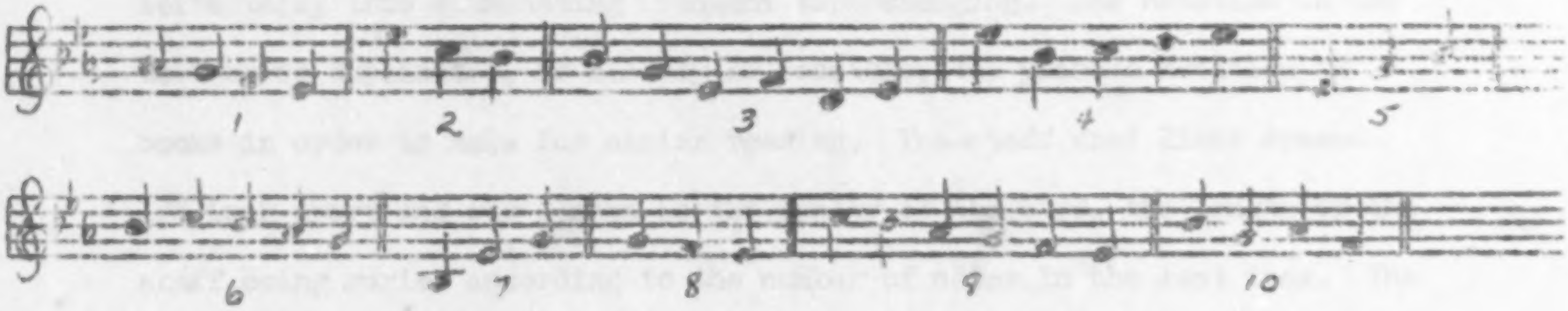
The second problem was concerned with the method to be used in presenting the items in the testing situations. The duration of the exposure time, while not a critical point in the study, needed to be uniform from individual to individual if the data were to be adequately interpreted. A memory drum having a six-second exposure time was considered a highly desirable method of maintaining a constant exposure time. The 20-item test, randomized for each of three successive trials, with each item duplicated in order to make possible two consecutive visual presentations, was copied in notation form and placed on the memory drum. A tape recording containing the starting pitch of each item, followed in six seconds by the complete aural presentation of the item, was synchronized with the rotation of the memory drum. The piano was used in preference to other media because the children were familiar with its sound. The test was then administered to a second exploratory sample of 10 fifth and sixth grade children. The results showed the six-second exposure time to be too short for an adequate response and the procedures again had to be revised.

Since the exposure time for the memory drum could not be increased without making substantial mechanical alterations in the mechanism, the writer placed all of the items on 4 x 6 white index cards. A second tape-recording was made containing the starting pitch of each item, followed in ten seconds by the complete aural presentation of the item. The examiner was to hand the card to the subject upon verbal cue from the tape and the subject retained the card until a new card was given to him. A third exploratory group of seven fifth grade children was given this revised test. The results showed the exposure time to be adequate and that fatigue did not appear to affect overall test performance. However, the three trials of the test took approximately 22 minutes which, when added to the time necessary for orientation and instruction, was considered too long. The time element was an important factor in working out the final form of the test since arrangements to use children had been approved with the limitation that all testing was to be done during the child's regular music period. The writer also felt that a shorter test would be desirable because the children should not be expected to concentrate on a test of this nature for much longer than fifteen minutes.

The final form of the test contained three trials of ten items and used four three-note items, two four-note items, two five-note items, and two six-note items. The items were randomized for each successive trial to eliminate the possibility of rote and serial learning. One trial of the final form of the test is presented in Figure 2.

Figure 2

(The 10-Item Test for Phase I)



5. Testing Procedures

The subjects for Phase I of the study were selected at random from grades four, five and six as previously described. All subjects were tested individually in a room other than a classroom. Prior to taking the test each subject was given a brief orientation which included a statement of the purpose of the experiment together with descriptions of the function of the equipment and of the tasks to be performed. The first trial of the 10-item test included a five-item practice trial which utilized the regular test procedures and gave the subject an opportunity to ask questions about anything that was not understood. Considerable emphasis was given to the idea that hearing an item after the subject had tried to read it was designed to make it possible for the subject to recall the sound of the item the next time it was presented visually, as well as to provide a model against which he could compare the original attempt.

Equipment used in the test situation included: (1) two tape recorders (Revere, Model T-700D), one for playing the master tape containing the aural parts of the test and the second for recording all pupil responses; and (2) three complete sets of test cards with ten cards to each set. Three sets made it unnecessary to re-sort the cards after each trial. The pupil responses were recorded at a tape speed of 3.75 inches per second. It had previously been established that this speed was entirely satisfactory for collecting accurate data and the slower speed also made it possible to conserve tape, thus eliminating frequent tape-changing. The notation on the test cards duplicated, as closely as possible, the printed notation in books in order to make for easier reading. The staff used lines spaced 1/8 inch apart and was placed in the center of the card, the length of the staff being varied according to the number of notes in the test item. The note heads were oval shaped with a lengthwise diameter of 1/8 inch and were spaced 3/8 inch apart. The treble clef sign, flats, and quarter-note stems were also of suitable size. No subject reported difficulty in reading the notation so it may be assumed that the visual stimuli were satisfactorily notated.

The testing of the first fifteen subjects revealed no significant weaknesses in either the test itself or the procedures being followed. The subjects were all interested in the test, particularily in hearing a short section of their recorded responses, and were enthusiastic and cooperative. Both the subjects and the examiner did feel that the ten-second interval was too long for some of the shorter items, particularilly between the aural presentation of an item and the visual presentation of the next item. A second master tape was made and the time intervals were varied depending on the length of the items, a change, which seemed to improve the "pace" of the entire test. Timing of the tape was managed by setting an electric metronome so it marked every second (60 beats per minute), using the flashing light rather than the clicks to keep the tape free of unnecessary sounds. The examiner gave the verbal cue for handing the stimulus card to the subject by saying "Card number __" within a two-second interval. The starting pitch was sounded on the third second and sustained for two seconds. This was followed by an interval of silence during which time the subject was to sing a response to the visual presentation. A second time interval then followed which included first playing the configuration at a rate of two quarter notes per second, followed by silence during which the subject was to sing what had been played. The beginning of each item was always the same, "Card Number ___, (Sound pitch)", the "*" occurring each second. The varying time intervals are shown below:

Length of item	Duration for first interval of silence	Duration of second time interval
3 notes	8 seconds	5 seconds
4 notes	9 seconds	6 seconds
5 notes	10 seconds	7 seconds
6 notes	11 seconds	8 seconds

No other changes were made during the remainder of the testing in Phase I.

6. Analyzing and Processing the Data

Processing the data presented no significant problems. The tapes contained all the responses of the subjects and these responses were transferred to individual data sheets for each subject. If a response was questionable the tape was merely reversed, the response replayed, and the data verified. Correct responses for both the visual and aural stimulus were indicated by a check mark; failure to try an item or to sing only the initial tone were indicated by a dash; incorrect responses, even if only a single tone was wrong, were indicated by the tonic sol-fa syllables for the entire response. Items sung out-of-tune, either in part or completely, were counted as incorrect responses and the exact pitch reproduction so indicated. Although the items had been randomized for each trial, the data sheets were arranged so that Item I appeared at the top of each trial regardless of its actual location within that trial. This procedure simplified the analysis and tabulation of the data which was done later. For both the visual and aural responses, the scoring system used was the number of items for which correct responses were made in two of the three trials.

Phase II

1. Purposes

The purposes of Phase II of the study were: (1) to examine perception as the integrative, internalized process through which the individual organizes previous auditory and visual perception of given stimuli in order to react to these same stimuli as they appear in new learning situations; (2) to compare the music reading performances of children with average musical ability and children gifted musically; (3) to test the hypothesis that tonal configurations tend to retain their identity from one situation to another; and (4) to study the effectiveness of the research techniques being employed.

2. **General Plan**

Two kinds of learning situations were constructed which differed in sequence of tasks rather in in content. These will be designated as Sequence A and Sequence B. Each sequence contained two tasks and the subject was given approximately 15 minutes to complete each task. In both sequences the two tasks were separated by a 48-hour time interval. This time interval between tasks was used for two reasons: (1) the results of the testing in Phase I indicated that children of this age group were not able to give concentrated attention to tasks of this nature for sessions lasting longer than 15 minutes; and (2) the testing was to be done during the time the child met with the regular music class, these classes being scheduled on alternate days rather than daily.

These two sequences, together with their tasks may be diagrammed as follows:

	Task I	Task II
Sequence A	The subject attempts to learn five tonal configurations in eight trials following the procedures established in Phase I.	The subject attempts to learn a short song constructed from the five separate items presented in Task I.
Sequence B	The subject attempts to learn the short song constructed from the five test items.	The subject attempts to learn the five tonal configurations which had been presented in the song.

The material used in both sequences was the same and the testing procedures followed in each task were similar for both sequences. The major difference between the sequences was the order in which the tasks were presented to the subject.

Persons familiar with the procedures recommended in developing music reading skill will recognize that Sequence A is based on the assumption that preliminary learning of certain tonal configurations will result in more rapid learning of a song containing these configurations. Sequence B is based on the assumption that music reading skill should begin with the song and then proceed to the drill work on the musical elements contained in the song.

3. Construction of the Tests

The first step taken in Phase II was the construction of the tests for the two tasks. Analysis of the data from Phase I showed that it would be desirable to limit the testing time for each task to a fifteen-minute period in order to obtain maximum effort from the subjects. One could reasonably assume that more trials on shorter tasks would result in a higher percentage of success than fewer trials on longer tasks, especially since the writer hoped that all subjects would be able to learn the material. A test of five tonal configurations was then constructed on the basis of the data from Phase I.

The procedures used in the item analysis involved the following steps:

1. Analysis of variance had shown that there were no significant differences between the mean achievement scores of the fourth, fifth and sixth grade subjects studied in Phase I. Therefore these three grades were combined in carrying forward the item analysis.

2. The two groups of subjects were then identified; those 27 per cent making the highest test scores and the 27 per cent making the lowest test scores. There were 24 subjects in each group.

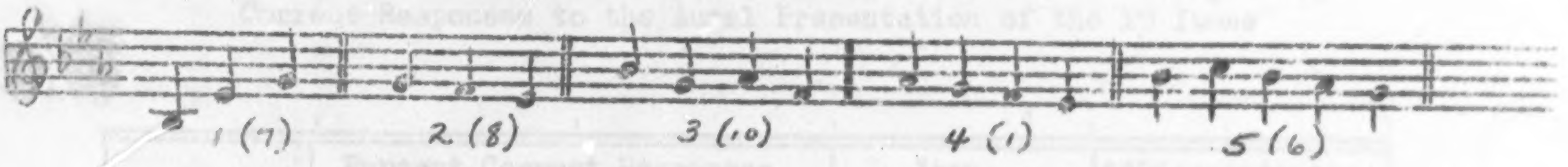
3. Item difficulty was computed by first determining the percentage of pupils in the upper group who made correct responses to the item and the percentage of pupils in the lower group who made correct responses. The average of these two percentages was defined as the difficulty value of the item. An item was scored with a correct re-

sponse if the subject made correct responses to the item on any two of three trials.

4. The discrimination index, or item-test correlation, was obtained by referring to a table suggested by Flanagan. (5)

The data for the item analysis, based on the percentage of correct responses for both the visual and aural aspects of the test, is presented in Table 10 and Table 11. Item difficulty was chosen as the criterion for selection rather than the discrimination index, although the later was also considered. It should be noted that the order of item difficulty varies, depending on whether the responses to visual or to aural stimuli are used as the basis for making the analysis. In view of the nature of the study the writer decided that the item analysis of visual responses should be considered in selecting the tonal configurations to be used and on this basis items 1, 6, 7, 8 and 10 were selected for use in Phase II. These items are shown in Figure 3.

Figure 3
The 5 Item Test for Phase II)



Although there were no sifnificant differences between the mean scores on the visual part of the test for the fourth, fifth and sixth grade pupils, the reader may be interested in comparing the results of the item analysis made for each grade. This data is presented in Table 12. It should be pointed out that the N for the upper and lower groups in each grade was quite small.

Table 10

Table Showing the Percentages of the Upper and Lower Groups Making Correct Responses to the Visual Presentation of the 10 Items

Item	Percent Correct Responses *Upper Group	 *Lower Group	Item Difficulty*	**Discrimination Index
1	92	21	57	70
2	21	4	13	53
3	17	0	9	--
4	42	4	23	68
5	96	42	69	77
6	83	8	46	70
7	42	8	25	40
8	92	33	63	63
9	100	50	75	72
10	63	4	34	77

*These figures are percentages.
Item Difficulty (Easy to Difficult): 9,5,8,1,6,10,7,4,2,3.
Discrimination Index (High to Low): 10,5,9,1,7,6,4,8,2,7,3.
**Flanagan (5)

Table 11

Table Showing the Percentages of the Upper and Lower Groups Making Correct Responses to the Aural Presentation of the 10 Items

Item	Percent Correct Responses *Upper Group	 *Lower Group	Item *Difficulty	**Discrimination Index
1	100	79	90	53
2	100	58	79	68
3	79	8	44	70
4	100	54	77	72
5	100	79	90	53
6	100	63	82	68
7	100	83	92	53
8	100	75	88	53
9	100	58	79	68
10	100	88	94	40

*These figures are percentages.
Item Difficulty (Easy to Difficult): 10,7,5,1,8,6,9,2,3,4.
Discrimination Index (High to Low): 4,3,6,9,2,1,5,7,8,10.
**Flanagan (5)

2

TABLE 12

Table Showing the Percentages of the Upper and Lower Groups in Each Grade Making Correct Responses to the Visual Presentation of the 10 Items

Item	4th GRADE			
	% Correct Responses		Item Difficulty	Discrimination Index
	Upper	Lower		
1	88	38	63	56
2	25	0	13	—
3	13	0	7	—
4	13	0	7	—
5	100	50	75	72
6	75	13	44	70
7	50	0	25	—
8	75	38	57	42
9	100	63	82	68
10	50	0	25	—
	Item Difficulty (Easy to Difficult) 9, 5, 1, 8, 6, 7, 10, 2, 3, 4			

Item	5th GRADE			
	% Correct Responses		Item Difficulty	Discrimination Index
	Upper	Lower		
1	100	0	50	—
2	29	14	22	30
3	14	0	7	—
4	57	0	29	—
5	86	43	65	56
6	86	14	50	77
7	43	0	22	—
8	100	29	65	81
9	100	43	72	77
10	57	0	29	—
Item Difficulty (Easy to Difficult) 9, 5, 8, 1, 6, 4, 10, 2, 7, 3				

TABLE 12 (Continued)

	6th GRADE			
	% Correct Responses		Item	Discrimination
Item	Upper	Lower	Difficulty	Index
1	89	22	56	70
2	11	0	6	—
3	22	0	11	--
4	67	11	39	63
5	100	44	72	77
6	89	0	45	--
7	44	22	33	24
8	100	33	67	81
9	100	44	72	77
10	78	11	45	70
	Item Difficulty (Easy to Difficult) 9, 5, 8, 1, 10, 6, 4, 7, 3, 2			

Since the tonal configurations were again to be presented as quarter notes without time signatures or measure bars, the song to be used should not contain significantly different notation. It was decided to write a song rather than attempt to find one in song-texts which could serve the purpose. These guides were used in constructing the song:

1. It should be reasonably musical without duplicating any familiar songs.

2. Two phrases, each containing four measures, was considered an ideal length.

3. No material other than the five tonal configurations already selected was to be included. This would, however, permit the repetition of at least one tonal configuration or of single tones.

4. Notation was to use only quarter and half notes, with each tonal configuration appearing as a single measure whenever rhythmically pos-

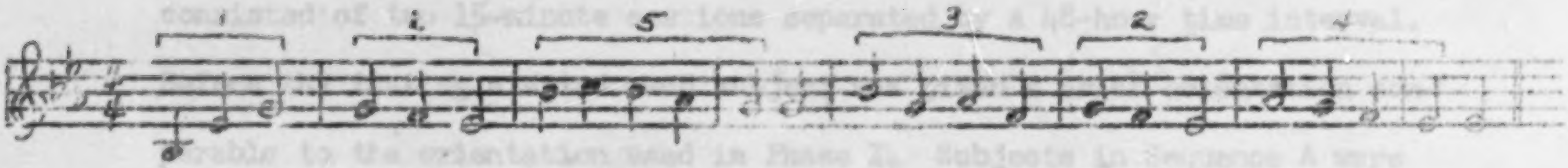
sible.

5. The song would be in the same key as the tonal configurations and would encompass a reasonable singing range.

Ten different songs were constructed and submitted to an informal jury of composition students, theory teachers and public school music teachers for reactions and comments. There was considerable agreement that two of the 10 songs were musically more interesting than the others and the test song was then selected by tossing a coin. This song, with the tonal configurations identified, is presented in Figure 4.

Figure 4

(The Song for Phase II)



4. Testing Procedures

The subjects for Phase II of the study were selected as described earlier in the report. Each of the three schools had four groups of subjects: (1) fourth grade, average; (2) fourth grade, gifted; (3) sixth grade, average; and (4) sixth grade, gifted. The subjects within each of these four groups were then assigned at random to either Sequence A or Sequence B, an attempt being made to have equal numbers of subjects in each of the sub-groups. This assignment to sequences gave eight groups for each school as shown in Table 13.

TABLE 13

Table Showing the Initial Assignment of
the 151 Subjects Participating in Phase II

	4th Grade				6th Grade				Totals
	Sequence "A"		Sequence "B"		Sequence "A"		Sequence "B"		
	Average	Gifted	Average	Gifted	Average	Gifted	Average	Gifted	
School A	10	5	10	5	11	8	10	7	66
School B	7	4	8	3	4	4	4	4	38
School C	7	5	7	5	7	6	5	5	47
Totals	24	14	25	13	22	18	19	16	151

Each subject was tested individually, as in Phase I, and the testing consisted of two 15-minute sessions separated by a 48-hour time interval. Before the testing started each subject was given a brief orientation comparable to the orientation used in Phase I. Subjects in Sequence A were told that during the first session they would try to learn to read five tone groups, or "musical words", and that in two days they would try to learn to read a song which used the same tone groups. Subjects in Sequence B were told that during the first session they would try to learn to read a song and that two days later they were to learn to read five tone groups, or "musical words" which had been used in the song. The orientation also included a short practice session in reading either four tonal configurations or two measures of a song, depending upon the sequence the subject was following.

The procedures for the two tasks were the same regardless of the sequence being followed. The process for learning the five items was identical with the process followed in Phase I and described in detail on pages 40, 41, and 42. There were eight trials of the five items, with the first trial considered as a pre-test, and the items randomized for each successive presentation. Learning the song followed the procedures outlined below:

1. The subject, with the notation of the song before him, is given the pitch of the first note of the song and asked to sing as much of it as possible. This served as a pre-test and was used as a basis for determining the amount of learning which took place during the testing.
2. This trial performance was followed by having the subject listen to a performance of the entire song, played at a rate of two quarter notes per second. During the listening the subject was encouraged to follow the notation and, if desired, to sing along with the piano.
3. Following this listening, the subject was asked to sing through the song a second time.
4. Steps 1 and 2 were repeated until the subject had had an opportunity to complete at least eight trials of the song, a trial being defined to include both the singing and listening aspects.

The reader may infer, from these procedures, that a process of rote learning might well be taking place rather than a process of learning to read the notation of the song. To ascertain whether this actually occurred a check was made of every fourth subject tested. At the conclusion of the testing session the notation of the song was removed and the subject asked to sing the song from memory. Only a very few subjects were able to proceed as far as the end of the fourth measure; none were able to sing farther than that point; and most subjects were unable to sing more than the first measure and part of the second. These same subjects were, however, able to proceed much farther when using the notation and so it appears reasonable to assume that actual reading rather than rote learning was resulting from the testing.

One significant difference existed between the procedures followed in learning the patterns and in learning the song; each pattern was presented individually for both the reading and listening attempts while the song was always presented as a complete entity. In order to minimize this difference, the writer decided that following each reading trial of the song, and prior to the subsequent listening, the reading errors made by the subject would be pointed out and verbally described by the examiner. This decision was based on the analysis of the data from Phase I which showed that many subjects persisted in making the same errors because of an inability to retain the sound of their reading attempt as a basis for comparison with the item when it was played for them. If this held true for items treated individually, it would be more true when the subject was expected to retain the aural memory of an eight-measure song. It was also felt that this assistance resulted in a better attitude on the part of the subjects because they were able to achieve better performances.

For both tasks a criterion of success had been established which consisted of two consecutive perfect performances, or trials, of the task. There seemed no valid reason to continue through all eight trials with subjects who could perform the task after two or three trials.

The equipment used in the test situations was the same as for Phase I and included: (1) two tape-recorders; (2) one set of test cards for the five items, there being sufficient time to re-arrange them for the next trial; and (3) one 5 x 8 index card containing the song, with the notation the same size as used in Phase I and described on page 50. The master tape for the five items was made according to the procedures outlined for the master tape in Phase I, pages 39 and 40; the master tape for the song contained ten successive performances of the song, although eight had been considered sufficient.

5. Analyzing and Processing the Data

Processing the data was carried forward in the same way as for Phase I. The data for both the patterns and the song were transferred to the individual data sheets and scores, in terms of correct responses, were calculated for each subject. Three kinds of scores were used:

1. A "per cent correct" score was calculated for each item in the five-item test and for each item in the song, the item being defined as the tonal configuration:

$$\text{"Per cent correct"} = \frac{\text{Number of trials item is correct}}{\text{Total number of trials}}$$

2. A "mean per cent correct" score was calculated for each task. This was a simple mean score:

$$\text{"mean per cent correct"} = \frac{\text{Sum of the "\% correct scores" on each item"}}{\text{Number of items in task}}$$

For both of these scores the unit which served as the basis of the score was the individual item. The first trial was not included in the total number of trials, as explained later.

3. A "rate of learning" score was calculated for each task and represented an attempt to locate a single score which could describe the true net gains made on the task and, at the same time, take into consideration the number of trials needed to make those gains. The formula used for this "RL score" was:

$$\text{RL} = \text{true net gain} \times \frac{100}{\text{Total N of trials} - 1}$$

The true net gain can be expressed as: $\frac{a-b}{c-b}$ in which:

a = Score on final trial, the final trial identified as the first of the two consecutively perfect trials.

b = Score on first trial.

c = Perfect score, the score assigned to any trial in which all items are correct.

The actual gain made by a subject is expressed in the numerator, "a-b"; the total gain possible for the subject is expressed in the denominator, "c-b". The result, which can never exceed 1, is the true net gain.

Since this true net gain does not reflect the number of learning trials taken by the subject, the second part of the formula was used. The numerator, 100, is used so the score may be expressed in larger units. The denominator indicates the actual number of trials during which learning could take place. The first trial for both tasks was considered a pre-test of previous knowledge, since it was assumed that learning the content of the task did not begin until the subject had heard that content. In the case of the song, this occurred after the initial trial; in the case of the tonal configurations, during the first complete trial but always after the individual item had been attempted. Therefore, this first trial was subtracted from the total number of trials, keeping in mind the definition of the final trial, to give the number of learning trials the subject had.

In Phase I the scoring of the items was done on a "right" or "wrong" basis. It was decided, in view of the nature of the tasks, to use a third alternative for scoring the items in Phase II. This involved giving credit for one-half a correct response if the subject sang the item at the wrong pitch level but retained the same pattern and relative pitches of the original item. For example, if the subject, in reading the item "fa-mi-re-do", did not correctly locate the starting pitch and sang "sol-fa-mi-re" instead, he was given credit for one-half a correct response. This scoring procedure gave data which more accurately described the music reading abilities of the subject since he was not penalized for his inability to sing independently. There are a great many pupils in the elementary grades who can read and interpret music symbols with a high degree of accuracy but who need the support of an accompanying instrument to maintain perfect tonality and intonation. The writer did not, in this exploratory study, wish to discriminate against the child for whom the task was already very difficult.

In order to illustrate the entire scoring procedure, a sample data sheet used for scoring the song is given in Figure 5 and the three scores are indicated, as well as the process of deriving the scores. The tonal configurations appearing in the song are indicated with brackets and the individual notes outside the bracket in measures 4 and 8 are not included in the scoring. A value of eight points is given for each correct tonal configuration yielding a total perfect score of 48 for the six items.

The data for Phase II was analyzed for each of the eight sub-groups in each of the three schools. Comparisons between groups were made in terms of the RL scores and the "per cent correct" scores. These groups are as follows:

1. 4th Grade, average pupils, Sequence A
2. 4th Grade, average pupils, Sequence B
3. 4th Grade, gifted pupils, Sequence A
4. 4th Grade, gifted pupils, Sequence B
5. 6th Grade, average pupils, Sequence A
6. 6th Grade, average pupils, Sequence B
7. 6th Grade, gifted pupils, Sequence A
8. 6th Grade, gifted pupils, Sequence B

Figure 5

DATA SHEET, PHASE II, SONG

Mary Jones (Name) A (Sequence) Gifted (Group) A-6 (School-and Grade)

	Item 1	Item 2***	Item 5		Item 3	Item 2***	Item 4		
	1	2	3	4	5	6	7	8	
	S D M	M R D	S L S F	M M	S M F R	M R D	F M R	D D	
1	✓	m d s	s s m d	d d	r t r t	r t s	m r d	s s	8
2	✓	m d s	s l s m	d d	r d r t	r t s	*s f m	r d	12
3	✓	✓	✓	✓	✓	s m d	*s f m	r d	36
4	✓	✓	✓	✓	✓	s m d	*s f m	r d	36
5	✓	✓	✓	✓	✓	✓	f m d	d d	40
** 6	✓	✓	✓	✓	✓	✓	✓	✓	48
7	✓	✓	✓	✓	✓	✓	✓	✓	
8									
1	5/5=100%	4/5=80%	4/5=80%		4/5=80%	2/5=40%	2.5/5=50%		

1=% correct score for each item.

*Item is counted as half correct and given 4 points because the intervallic relationship is correct even though the pitches are incorrect.

**Identified as final trial because it is the first of two consecutively perfect trials.

***This item is repeated and accounts for the six items of the task.

Mean % correct score= $\frac{100+80+80+80+40+50}{6}$

$= \frac{430}{6} = 71.6$

RL Score = $\frac{48-8}{48-8} \times \frac{100}{5} = \frac{40}{40} \times \frac{100}{5} = 20$

RESULTS OF THE STUDY, PHASE I

Usable data was obtained for 89 of the more than 100 pupils tested in Phase I. Since the subjects had been selected at random from the total four fifth and sixth grade population in each school it was highly possible that some subjects would be tested whose data would not be usable. Such subjects would include the following:

(1) pupils who verbally expressed no desire to be tested;

(2) pupils who were unable to match pitches and consequently would not be able to make any singing response to either the visual or aural stimulus.

(3) pupils whose ability to repeat what they heard was so poor that they were unable to sing more than two correct responses out of the ten items;

(4) pupils who became so nervous during the testing that they were unable to continue.

In discussing the data the term "visual" will be used to refer to the responses made to the visual presentation of the items--the reading aspect of the test. The term "aural" will be used to refer to the responses made to the aural presentation of the items--the listening aspect of the test.

Visual Part of Phase I

The data for pupils in each grade in each school were first tabulated to ascertain the kinds of scores which had been made. The basis of scoring was to count an item correct if a correct response had been made for any two of the three trials. An item was scored wrong if it was incorrectly sung or if no trial or only a partial trial was made of the item. The score was the sum of the individual items correct, with a maximum score of 10.

Table 14 summarizes the mean scores for each grade in each school.

TABLE 14

Table Showing the Mean Scores for the Visual Part of Phase I for each Grade in each School

School	4th Grade		5th Grade		6th Grade		Combined Grades	
	N	Mean	N	Mean	N	Mean	N	Mean
A	10	4.2	10	4.9	11	4.36	31	4.5
B	9	2.9	9	2.3	10	4.1	28	3.1
C	11	3.5	9	4.9	10	5.1	30	4.4
Total	30	3.53	28	4.07	31	4.52	89	4.0

The data shows the mean score for the fourth grade to be lower in each school, than the mean scores for the two upper grades. The same difference appears when the three schools are combined by grade level. The data also shows mean score for the sixth grade to be higher than the mean score for the fifth grade in Schools B and C, but not in School A. When the three schools are combined, however, there is a steady increase in the size of the mean from fourth grade through sixth grade.

A test for the significance of the differences between the mean scores of each grade, combining the three schools, gave F= 1.5068, and it may be concluded that there is no significant difference between the three grades in terms of the data of Phase I. The five per cent level of significance is used throughout the study. This, of course, has significant implications if intensive testing at these levels continued to support these findings. A test for the significance of the differences beteen the scores for the fourth and sixth grades, the grades to be studied in Phase II, gave a "t"

TABLE 15

Analysis of Variance between Grades for the Visual Part of Phase I

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
Total	435.8202	88	
Grades	14.7545	2	7.3773
Pupils in Grades	421.0657	86	4.8961
			(F=1.5068)

"Distribution of F"--Johnson-Palmer, *Statistical Methods in Research*, (New York: Prentice-Hall, 1949 pp. 362-365)

of 1.8397 and the differences were considered as not significant.

$$"t" = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\left(\frac{\Sigma x_1^2 + \Sigma x_2^2}{N_1 + N_2 - 2}\right)\left(\frac{1}{N_1} + \frac{1}{N_2}\right)}} \quad \text{in which } \Sigma x^2 = \frac{1}{N}\left[N\Sigma X^2 - (\Sigma X)^2\right]$$

Table 14 also shows the mean scores for each grade to vary from school to school. For the fourth grade School A had the highest mean, with schools C and B in descending order. For the fifth grade, schools A and C had identical means and for the sixth grade School C had the higher mean, with schools A and C in descending order. This would appear to indicate that subjects in each grade in School B, consistently performed at a lower level of achievement than subjects from the other two schools. The mean scores for the combined grades in each school shows schools A and C to have overall means which are very similar and, at the same time, more than one point higher than the mean for School B.

Analysis of variance yielded an F= 3.5553 which is significant at the 5% level and one may say that a significant difference exists between the three schools. Further analyses showed the following:

TABLE 16

Analysis of Variance Between Schools
For the Visual Part of Phase I

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares
Total	435.8202	88	
Schools	33.2830	2	16.6415
Pupils in Schools	402.5372	86	4.6807
			(F=3.5553)

1. There is no sifnificant difference between the mean scores for the combined grades of School A and School C; "t"= .0907.

2. There is a sifnificant difference between the mean scores of School A and School B; "t"= 2.394.

3. There is a significant difference between the mean scores of School C and School B; t= 2.2249.

This analysis shows that the subjects from School B did not perform at as high a level as did the subjects from either of the other two schools. In order to accurately identify the reasons for this difference, it would be necessary to consider such factors as: (1) the socio-economic background of the subjects; (2) the nature of the music program in this school as compared to the other schools; (3) the music teacher; and (4) the musical interests and music training of the subjects as compared to subjects in other schools.

The teacher rating sheets, to be discussed later in the report, showed that the pupils who were placed in the lowest group in terms of musical ability, were always the boys. This was true of the ratings made by each teacher in each school. The professional literature abounds with articles which consider the "boy problem" in music activities at the various grade levels. There is however, nothing to suggest that the objectives of the music program differentiates between boys and girls and there is no information to indicate that girls possess more native musical ability than do the boys. The conclusion which must be reached is that the girls have a different attitude toward music which makes for better achievement, more interest, and consequently for higher teacher ratings.

The subjects for Phase I had been selected in such a way so as to secure equal numbers of boys and girls at each grade in each school. A summary of the results on the visual part of the test is given in Table 17. This shows

TABLE 17

Table Showing the Mean Scores for the
43 Boys and 46 Girls Tested in Phase I

	4				5				6				Combined			
	B		G		B		G		B		G		B		G	
School	N	Mn	N	Mn	N	Mn	N	Mn	N	Mn	N	Mn	N	Mn	N	Mn
A	5	4.6	5	3.8	5	4.8	5	5	5	4.6	6	4.2	15	4.7	16	4.3
B	4	2.5	5	3.2	4	2.5	5	2.2	5	3.4	5	4.8	13	2.9	15	3.4
C	5	3.2	6	3.7	5	3.0	4	7.2	5	4.0	5	6.2	15	3.4	15	5.5
Totals	14	3.5	16	3.6	14	3.5	14	4.6	15	4.0	16	5.0	43	3.67	46	4.39

the following:

1. For School A, the boys earned higher scores than the girls in both the fourth and sixth grade and these scores were very close together in fifth grade. When the scores for three grades were combined the boys had a higher mean than did the girls.

2. For School B, the girls earned higher scores than the boys in both

the fourth and sixth grades. When the scores for the three grades were combined the girls had a higher mean than did the boys.

3. For School C, the girls consistently did better than the boys.

4. For the three schools combined, only the fourth grade showed little difference between the boys and the girls.

5. For the combined grades in the combined schools the boys had a mean of 3.67 and the girls had a mean of 4.39.

The "t" test of significance gave a "t" of 1.5457 and it may be concluded that there is no significant difference between the mean scores of the boys and girls when all grades and all schools are combined.

Analysis of Reading Errors for the Visual Part of Phase I

A considerable amount of time was spent in making a detailed analysis of the reading errors made on each of the ten items of the test. The purposes of the analysis were: (1) to identify any common errors which occurred; (2) to attempt to identify the progress shown by the subject for the three trials; and (3) to provide necessary data for the item-analysis. The data on the item analysis has already been presented in Section V and will not be reviewed here. Within the scope of this study there has not been sufficient time to make the complete analysis of data necessary for the second purpose cited, although the data is tabulated and available. This section will, therefore, present the findings of the analysis of errors only.

In making the analysis the responses to each item were considered separately for each trial so that a perfect score on all 10 items was 30 points. Multiplying 30 by the number of subjects in each grade yields the total number of responses possible. The analysis classified three types of responses: (1) right; (2) wrong; and (3) no trial. A "no trial" consisted either of no attempt to sing a response or failure to sing more than one or two notes of the item.

The data presented in Table 18 summarizes the responses of the 89 subjects, given by grade. Inspection of the totals and means in the table shows the following:

1. Fourth grade pupils had twice as many "no trial" responses as either of the other two grades, with a mean of 2.5. Fifth and sixth grade "no trial" means are very similar.
2. Sixth grade pupils had a lower mean of "wrong" responses than did the other two grades; 15.8 compared to 16.4 and 16.8 for fourth and sixth grades.
3. The mean "right" responses for the three grades show improvement from grade to grade although it is not pronounced.

Consideration of the data for each item will be considered as each item is discussed individually.

The scope of this report is such that a detailed discussion of each item would prove impractical. The writer proposes, therefore, to give only the following information:

1. The number of "no trial", "wrong", and "right" responses to the item, by grade and by combining grades, for the three trials combined.
2. A summary of the most frequently made errors by the combined grades, using the tonic sol-fa syllables for identification rather than actual notation.

Table summarizing the Responses made by 89
Subjects to the Visual Part of Phase I.

Item	4th Grade				5th Grade				6th Grade			
	N	No Trial	Wrong	Right	N	No Trial	Wrong	Right	N	No Trial	Wrong	Right
1	30	10	32	48	28	3	33	48	31	4	33	56
2	30	8	69	13	28	7	62	15	31	4	80	9
3	30	15	69	6	28	10	69	5	31	6	80	7
4	30	14	62	14	28	4	57	23	31	3	57	33
5	30	4	27	59	28	0	31	53	31	4	27	62
6	30	6	55	29	28	3	50	31	31	3	47	43
7	30	4	63	23	28	3	62	19	31	3	61	29
8	30	4	34	52	28	4	29	51	31	2	25	66
9	30	4	17	69	28	2	18	64	31	2	23	68
10	30	6	63	21	28	1	61	22	31	1	58	34
Totals	30	75	491	334	28	37	472	331	31	32	491	407
Mean		2.5	16.4	11.1		1.3	16.8	11.8		1.0	15.8	13.1

Item 1: (F M R D)*

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	10	3	4	17
"Wrong"	32	33	33	98
"Right"	48	48	56	152

Within the 98 errors made by the combined grades, the following appeared most frequently:

1.	FRTS	10	4.	[illegible]	5	7.	FRDTe	4
2.	FRTF	5	5.	FSLT	5	8.	SMDS	5
3.	FRMD	5	6.	FMeRD	4	9.	SPMR	6
							Total	49

The remaining 49 errors had frequency counts of 1 to 3. 25 errors had a frequency count of only 1. It is interesting to note that Error [illegible] is an exact inversion of the pattern and Error 4 a partial inversion; that errors 1, 2, and 8 are descending chord figures; and that Errors 8 and 9 began on the wrong pitch despite the fact that the actual *starting pitch was sounded for two seconds. In fact, 27 of the* 98 errors included failure to begin on the correct pitch. The scores for the three grades are very similar although the fourth grade had more "no trial" responses.

Item 2: (D L E)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	8	7	4	19
"Wrong"	69	62	80	211
"Right"	13	15	9	37

Within the 211 errors of the combined grades, these appeared most frequently:

1.	DTL	71	4.	DSL		7.	DTE	5

2.	DSM	44	5.	DSF	9	8. DTeL	4
3.	DLF	26	6.	DSD	8	Total	177

The eight **errors** account for 177, or 64 per cent of the total wrong responses. Of remaining 34 errors, 17 had a frequency of 1, four and a frequency of 2, and three had a frequency of 3. It is interesting to note that Errors 1, 2 and 3 were the most common. Children do not appear to notice the interval of the minor third followed by the major second since they sing the pattern either as a partial scale or as a complete chord figure. There were only 37 correct responses out of a possible total of 267. It is of interest to note that the sixth grade had more "wrong" responses and fewer "right" responses on this item than either of the other two grades. Only 17 errors were caused by failure to start on the correct pitch.

Item 3: (S M D $\bar{R}$ $\bar{T}$ $\bar{D}$)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	15	10	6	31
"Wrong" /	69	69	80	218
"Right"	6	5	7	18

The errors appearing most frequently in the total of 218 were:

1.	[illegible]	14	4.	[illegible]	10	7. [illegible]	6
2.	[illegible]	13	5.	[illegible]	7	8. [illegible]	5
3.	[illegible]	13	6.	[illegible]	6	9. [illegible]	5
						Total	79

These 9 errors account for only 79, or 3.6 per cent of all errors with **most of the** remaining 139 having frequency counts of only 1 or 2. This item was the most complex because it included six tones and was very disjunct, and this accounts for the fact that it had

the fewest number of correct responses of any item. The same h
true for the aural responses to the item, it again had the lowe
number of correct responses.

Errors 1, 7, 8 and 9 show a partially correct response being made, presumably because the first part of the pattern was famil and could be remembered better than the last three notes. Error 5 is a simplification of the pattern so it becomes a scale while t other errors show a tendency to repeat a 3-tone pattern. Failur to begin on the starting pitch accounts for only 32 of the error.

Item 4: (D $\underline{S}$. $\overline{L}$ $\overline{T}$ $\overline{D}$)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	14	4	3	21
"Wrong"	62	57	57	176
"Right"	14	23	33	70

This item began with a descending interval of a fourth and cont with an ascending scale pattern which returned to the starting to The interval is a difficult one to sing and this is evident in t number of errors. The 9 errors most frequently made accounted f 100, or 57 per cent of the total "wrong" responses.

1.	DTLSF	32	4.	DDRMF	13	7.	DFSLT	5
2.	DTLTD	19	5.	DMFSL	8		Total	100
3.	DLTDR	16	6.	DLSFM	7			

Errors 1 and 2 completely ignore the beginning interval in the item, either because it was not perceived or because the subjects were unable to produce any interval larger than a second. Error 6 underestimates the size of the beginning interval but does acknowledge its presence, and then ignores the ascending scale and transforms it into a descending scale. The remaining errors all show recognition of the general shape of the item, a skip down followed by an ascending scale, but fail to get the interval correct. This characteristic was also true for most of the remaining 76 errors which had frequencies of 1, 2 or 3. The sixth grade had considerably more "right" responses to the item than the other two grades. Only 13 errors were the result of an inability to match the starting tone which was given.

Item 5: (D $\overline{M}$ $\overline{S}$)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	4	0	0	8
"Wrong"	27	31	27	85
"Right"	59	53	62	174

This is one of the more familiar items on the test, a fact reflected in the large number of "right" responses by all grades. Note the way in which the pupils in the fourth and sixth grades duplicate each others performance. This item had only 85 "wrong" responses, a record exceeded only by item 9, and one error, "D R M", had a frequency of 53, or 62 per cent of the total. In fact, there were only 18 different errors counted for this item, the other 17 having frequencies of 1 or 2. The tonic chord is often sung as a 3-note scale while DRM is often sung as the tonic chord. This suggests that subjects do not really read the notation but more often guess at it. A total of 13 errors resulted from failure to begin on the correct pitch.

Item 6: (S L S F M)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	6	3	3	12
"Wrong"	55	50	47	152
"Right"	29	31	43	103

This combination of neighboring tones and a partial descending scale was easier for the sixth grade than for fourth and fifth, whose scores are similar. A total of 62 different errors were counted and the 7 most frequently made accounted for 75, or 49 per cent of the total number of "wrong" responses. These are:

1.	SLSMD	20	4.	SMFRD	9	7.	SLSMR	5
2.	SFMRD	15	5.	SLFMR	8			
3.	SDSMD	12	6.	SLSFR	6		Total	75

Errors 1, 6 and 7 show a partial success in reading the item, while Error 2 results again from attempting to use a familiar response to notation which is not carefully observed. Errors 3, 4 and 5 represent inaccurate responses which were strongly influenced by the aural presentation of the item in the trial which preceded this particular item. Only 12 errors were identified as failure to begin on the correct note.

Item 7: (S D M)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No Trial"	4	3	3	10
"Wrong"	63	62	61	186
"Right"	23	19	29	71

This item, an inversion of the more familiar item 5, contains within its first two notes one of the most common tone patterns identified in songs.

It ranked first in the analysis made by the writer and Smiley's analysis and is invariably found in songs which have an anacrusis, or "pick-up note." Despite this, the fact that it begins with an ascending fourth presents a definite reading problem to children in all the grades. Many children experienced difficulty in singing the first tone because they were unable to relax sufficiently for the note to sound. This difficulty did not appear in the analysis of aural responses to the same item although for the reading response it apparently was a factor since 47, or 25 per cent of the total errors were the result of an incorrect beginning pitch. The common errors identified for this item are:

1.	$\underline{S}$TR	49	5.	STD	8	9.	DDM	4
2.	$\underline{S}$LT	36	6.	SMF	6	10.	DRS	4
3.	DRM	10	7.	ST$\underline{S}$	5		Total	135
4.	STD	9	8.	TRF	4			

The 10 errors account for 73 per cent of the total "wrong" responses and Errors 1 and 2 account for 46 per cent of this same total. The tendency to reduce the interval of a fourth to an interval of a third and proceed to sing a tonic chord is evident in Error 1; the tendency to revert to a scale pattern whenever the reader is in doubt is shown in Error 2.

Item 8: (M $\underline{R}$ $\underline{D}$)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	4	4	2	10
"Wrong"	34	29	25	88
"Right"	52	51	66	169

This is one of four items for which the number of "right" responses exceeds the number of "wrong" responses. The fourth and fifth grade pupils perform almost the same way, with the sixth grade showing better results. Only 17 of the 88 errors resulted from beginning on the wrong pitch. The fas

of this pattern is established by the scores. The common errors we

1. MSD	24	3. SMD	6	5. MFR	5		
2. MFS	10	4. MSL	5	6. SFM	4		
				Total	54		

These errors account for 61 per cent of the total and it is interestin

note the frequency with which this simple scale pattern becomes a descen

tonic chord, Error 1 has a frequency of 24, or 27 per cent of the total

of errors. The reader is referred to the analysis of Item 5 which shows

parallel results.

Item 9: (L S F M R D)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	4	2	2	8
"Wrong"	17	18	23	58
"Right"	69	64	68	201

This six-note descending scale proved to be the easiest item of the ten

all grades appeared to perform with equal ability. There were no errors

a frequency greater than five although a total of 37 *different* errors we

counted. Failure to begin on the starting tone did account for 20, or 3

per cent of the total errors made.

Item 10: (S M F R)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	6	1	1	8
"Wrong"	63	61	58	182
"Right"	21	22	34	77

The disjunct character of this item made it more difficult than one would

expect since a sequence of descending thirds occurs quite often in songs.

Again very little difference exists between the fourth and fifth grades and the sixth grade achieves slightly better scores. A total of [illegible] different errors were tabulated and all but the four given below had frequencies of six or less:

1.	SFMR	27	3.	SMRD	14		
2.	SMSM	14	4.	SMRD	10	Total	65

Error 1 again represents this phenomenon of substituting a familiar tonal configuration for an unfamiliar one, possibly the result of carelessness but more often resulting from the subject's inability to "hear" the pattern before it is sung. Only 9 errors of the 182 were caused by an incorrect starting pitch.

Summary of the Findings

This detailed analysis of reading errors emphasizes several important points. The total of 1454 "wrong" and 114 "no trial" responses as compared with 1072 "right" responses shows that children in the elementary grades experience considerable difficulty in reading tonal configurations which are commonly found in the songs children sing.

1. During the testing period the writer found that most children read the item very slowly, a single note at a time. This is not apparent in the data because there is no way of indicating this kind of performance other than to listen to the tape-recorded responses.

2. Although children experienced little difficulty in making a response to the aural stimulus they were seemingly unable to "make the voice go the way the notes do", as several subjects pointed out. This implies that skill in reading music can be developed only if the child is able to "hear" the item silently before he sings it. Developing this internal "hearing" presents many problems including an adequate identification of

effective teaching procedures.

3. Children are usually aware of the general shape or direction of the tonal configurations but fail to acquire the ability to accurately identify the internal changes which make one pattern different from another even though the two patterns appear quite similar.

4. When children are asked to read tonal configurations which are not immediately recognized as something they know, one of two types of responses generally results: (1) changing the pattern so that it becomes a familiar pattern which is very similar to the stimulus, or (2) by guessing that the response being made is something like the stimulus without having any sound basis upon which to judge the accuracy of their response. The total number of different errors which the children made on the 10 items is evidence of the second kind of response. Excluding the "no trial" responses, 489 different errors were made in accumulating the 1454 "wrong" responses, giving an average frequency of 2.97 per error. Evidence of the first type of response is indicated by the fact that 61, or 12.47 per cent of the 489 different errors had a total frequency count of 767, or 53 per cent of the 1454 "wrong" responses.

5. The data from Phase I shows comparatively little difference between grades in terms of general ability to perform the task. Part of this might be due to the fact that three trials of the 10 items did not give the subjects sufficient opportunity to learn the items. The question arises as to whether more trials would have shown significant differences between the three grades. Since most music reading activities begin at

the third grade level. This means that sixth grade children have two additional years of experience with reading music than do fourth grade children. If there is no significant difference between the achievement of the two grades, and the data from Phase I shows that this is true in terms of this particular investigation, some highly interesting speculations may be made as to why this is true.

Test Scores on the Aural Part of Phase I

Paralleling the analysis of reading errors was an analysis of the errors made by the subjects as they responded to the aural presentation of each of the ten items. It is generally agreed by music educators that music reading ability depends a great deal upon developing the ability to perceive music aurally. This "ear vocabulary" of musical sounds establishes the basis for music reading activities because the child is then able to work with familiar concepts and materials.

The data on the singing responses to the aural presentation of the items was processed in much the same way as the data for the visual part of the study. The basis of scoring was again to count an item correct if a correct response had been made for any two of the three trials. An item was counted as wrong if it was incorrectly sung or if no trial or only a partial trial had been made on the item. The score was the sum of the individual items correct, with a maximum score of 10.

Table 19 summarizes the mean scores for each grade in each school. The data shows that fourth and fifth grade subjects, in each school and for the combined schools, had lower mean scores than the sixth grade subjects. Only

Table 19

Table Showing the Mean Scores for the Aural Part of Phase I for each Grade in Each School

School	4th		5th		6th		Combined	
	N	Mean	N	Mean	N	Mean	N	Mean
A	10	8.3	10	8.5	11	8.8	31	8.6
B	9	7.8	9	7.0	10	8.4	28	7.8
C	11	8.3	9	8.4	10	8.9	30	8.5
Total	30	8.2	28	8.0	31	8.7	89	8.3

TABLE 20

Analysis of Variance Between Grades for the Aural Part of Phase I

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
Total	219.1910	88	
Grades	8.0039	2	4.0020
Pupils in Grades	211.1971	86	2.4557
			(F=1.6299)

TABLE 21

Analysis of Variance Between Schools for the Aural Part of Phase I

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
Total	219.1910	88	
Schools	13.1195	2	6.5598
Pupils in Schools	206.0715	86	2.3962

In School B was there a marked difference between the means of fourth and fifth grade subjects. These results are similar to the results on the visual part of the test in terms of the relative achievement of each grade. The reader may wish to compare these means with the means on the visual part presented in Table 14, page 59. Such a comparison shows that pupils in each grade earned a score on the aural part of the test that was approximately four points higher than the scores for the visual part. In other words, they were able to double their achievement when an imitative, rather than a reading response was required.

A test for the significance of the differences between the mean scores of each grade, combining the three schools, gave F= 1.6297 and it may be concluded that there is no significant difference between the three grades in terms of this data. (See Table 20, page 75)

This may well indicate that the ability to respond to an aural stimulus is already developed by the fourth grade level and does not change significantly after that time.

Table 19 shows the mean scores for each grade in School A and C to be very similar but that subjects in School B consistently performed at a lower level of achievement, especially when the fifth grades are compared. Analysis of variance between the schools gave F= 2.7376 and it may be concluded that no significant difference exists between the three schools when subjects are asked to respond to an aural stimulus. (See Table 21, page 75)

There is no reason to suspect that boys differ significantly from girls in ability to respond to an aural stimulus except the difference which may be due to the general inclination of the boy to avoid using his singing voice.

The combined mean score for the 43 boys tested was 6.023 and for the girls it was 5.587. The "t" test of significance gave "t"= 1.0366 and it may be concluded that there is no significant difference between the mean scores of boys and girls.

These data would tend to support the point-of-view that auditory perception, as measured by a singing response to an aural stimulus, is already developed by the time a child reaches the fourth grade and does not significantly change as a result of the additional musical experiences in the upper elementary grades.

Analysis of Singing Errors for the Aural Part of Phase I

A detailed analysis of the singing errors made on each of the ten lines was carried forward following the procedures used in making the analysis of reading errors. The reader has observed the fact that the subjects had considerably higher mean scores on the aural part of the test than they did on the visual part. The writer was particularly interested in whether subjects with high scores on the aural part of the test had also obtained high scores on the visual part. The product-moment correlations were computed and resulted in the following:

1. Sixth grade, $r = .64$
2. Fifth grade, $r = .70$
3. Fourth grade, $r = .58$
4. Combined grades, $r = .64$

The writer is unable, at this time, to determine whether these could be considered high, medium, or low correlations in terms of the two skills being considered. It is of interest to note that fifth grade children are [illegible] more consistent than children in the other two grades, with the fourth grade not maintaining as high a straight-line "relationship. This is a problem which needs further investigation before any sound conclusions can be drawn.

The data presented in Table 22 summarizes the responses of the [illegible] subjects, these responses again being considered separately for each trial so that a perfect score on all ten items is 30. The table shows the following:

1. There were very few "no trial" responses, only 16 for the three grades, with the sixth grade having only one.
2. Sixth grade pupils had fewer "wrong" responses, a mean of 4.1, while fourth and fifth grades had means of 5.67 and 6.1 respectively.
3. The mean "right" responses shows the sixth grade to be slightly better than the other two grades.

Since the reader may be interested in comparing the visual and aural parts of the test, Table 23 presents a brief summary of the pertinent information. This clearly points out the comparative difficulty of the two parts of the test.

The writer does not feel that any specific purpose would be served by a detailed, item-by-item analysis of the singing errors made by the subjects. Inspection of Table 22 shows little difference between grades for the three kinds of responses made to each item. This data is summarized in Table 24 and the writer proposes to discuss only those items for which a significant number of errors were made; Items 2, 3, 4, 6 and 9.

Table [illegible]

Table Summarizing the Responses made by 89
Subjects to the Aural Part of Phase I.

Item	4th Grade				5th Grade				6th Grade			
	N	No Trial	Wrong	Right	N	No Trial	Wrong	Right	N	No Trial	Wrong	Right
1	30	0	8	82	28	1	13	70	31	0	5	88
2	30	2	8	80	28	1	21	62	31	0	13	80
3	30	1	60	29	28	2	55	27	31	0	56	37
4	30	1	16	73	28	2	14	68	31	0	10	83
5	30	1	6	83	28	1	3	80	31	0	2	91
6	30	1	15	74	28	0	17	67	31	1	9	83
7	30	0	10	80	28	0	7	77	31	0	9	84
8	30	0	4	86	28	0	8	76	31	0	4	89
9	30	2	28	60	28	0	25	59	31	0	15	78
10	30	0	15	75	28	0	7	77	31	0	3	90
Totals		8	170	722		7	170	663		1	126	803
Mean		.03	5.67	24.1		.03	6.1	23.7		.003	4.1	26.0

TABLE 23

Table Summarizing the Responses made by 89 Subjects
to Both the Visual and Aural Parts of Phase I

Grade	NO TRIAL				WRONG				RIGHT			
	Visual		Aural		Visual		Aural		Visual		Aural	
	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean	N	Mean
4th	75	2.5	8	.03	491	16.4	170	5.67	334	11.1	722	[illegible]
5th	37	1.3	7	.03	472	16.8	170	6.1	331	11.8	6[illegible]3	[illegible]
6th	32	1.0	1	.003	491	15.8	126	4.1	407	13.1	803	26.0
Combined	144	1.6	16	.02	1454	16.3	466	5.2	1072	12.0	2188	24.[illegible]

The total number of possible responses is 2670, obtained by multiplying the number of responses on the three trials (30) **by the number of subjects (89).**

TABLE 24

Table Summarizing the Responses made by 89 Subjects
to the Aural Part of Phase I (Combined Grades)

ITEM	NO TRIAL	WRONG	RIGHT
1	1	26	240
2	3	42	222
3	3	171	93
4	3	40	224
5	2	11	254
6	2	41	224
7	0	26	241
8	0	16	251
9	2	68	197
10	0	25	242
Totals	16	466	2188

The total number of possible responses is 2670, obtained by multiplying the number of responses on the three trials (30) **by the number of subjects (89).**

Item 2: (D L S)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No Trial"	2	1	0	3
"Wrong"	8	21	13	42
"Right"	80	62	80	222

Only one common error was identified, DTL, from the total of 11 different errors counted. This error was made 15 times and suggests, as do all the errors, failure to listen carefully. Inability to remember the sound of the item resulted in a return to the familiar scale pattern. The fifth grade seemed to have slightly more difficulty with the item than the other two grades.

Item 3: (S M D R T D)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	1	2	0	3
"Wrong"	60	55	56	171
"Right"	29	27	37	93

This complex item also proved the most difficult item of the ten when presented aurally. The fact that a total of 102 different errors accounts for the total of 171 wrong responses appears to be sufficient evidence that children could not retain an aural perception of so complex a tonal configuration. There were only two errors having a frequency count of more than five; SMFRMD was sung 17 times and SMDMTD was sung 9 times. In general, the errors all followed the shape of the original stimulus but failed to take into account the intervals within the pattern. Only 40 of the 171 errors made reproduced the first three tones of the pattern correctly.

Item 4: (D S̲ L̄ T̄ D̄)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	1	2	0	3
"Wrong"	16	14	10	40
"Right"	73	68	83	224

No common error was identified for this item and 27 different errors accounted for the total of 40 "wrong" responses. Two errors had a frequency of five: DT̲L̲T̄D̄ and DT̲L̲S̲F̲; representing inadequate aural perception of the shape of the item.

Item 6: (S L̄ S̲ F̲ M̲)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	1	0	1	2
"Wrong"	15	17	9	41
"Right"	74	67	83	224

Twenty-four different errors were noted for this item, with no error having a frequency greater than three. This would indicate that the "wrong" responses were probably the result of guessing. The fifth grade again fails to come up to the performance of the other two grades.

Item 9: (L S̲ F̲ M̲ R̲ D̲)

	4th Grade	5th Grade	6th Grade	Total for Combined Grades
"No trial"	2	0	0	2
"Wrong"	28	25	15	68
"Right"	60	59	78	197

This descending scale pattern, had been found to be the easiest item on the visual part of the test, in terms of the number of correct responses, and the would expect the same to be true when the item was presented aurally. The data shows that the second highest number of "wrong" responses were made to this item, the highest number being made for Item 3. Three common errors were identified for the item; each with a frequency count of 12: (1) [illegible]; (2) SFMRCT; and (3) SFMRD. The essential character of the item is **preserved** and, had the subject been able to retain the sound of the beginning pitch he would probably have responded correctly. Inspection of the raw data shows that 60 of the 68 "wrong" responses included failure to start on the correct note. This type of error is significant only for Item 9, the other items do not show so high a percentage of "wrong" response due to this factor. It is not reasonable to say that it is caused by the length or complexity of the item because the data for Item 3, of equal length and greater complexity, showed that only 23, or 1.5 per cent of the total "wrongs", included failure to start on the correct pitch. The writer suggests that the very simplicity and familiarity of the scale pattern gave the subjects a sense of over-confidence so they failed to pay attention to the starting point of the scale. Error 3 shows that some subjects remembered the final tone and shortened the pattern so it conformed to what had been retained.

The inability to locate the correct beginning pitch for items presented aurally is probably due to the fact that the subject has to retain this pitch during the entire aural presentation. For the visual part of the study, with the subject singing the starting pitch immediately after it had been given, failure to begin on the correct pitch accounts for 1.4 per cent of the total errors; for the aural part of the study it accounts for 39.3 per cent of the total errors.

Summary of the Findings

The analysis of singing errors which has been briefly presented shows that aural imitation of an item is considerably easier for children than making responses to visual stimuli. While 54.5 per cent of the responses to the visual presentation of the items were "wrong" responses, only 17.5 per cent of the responses to aural presentations were "wrong". This again emphasizes the difference in the difficulty of the two kinds of responses. The "wrong" responses to the aural presentation of the items showed few common errors and this suggests that guessing is taking place. It appears unlikely that music reading skill can be developed by continuing to emphasize rote-learning and imitative singing in the upper grades. The data shows that little difference exists between grades in terms of this auditory perception. The major problem centers around the methods used to help the child relate this auditory perception to the visual stimuli so the latter are meaningful.

RESULTS OF THE STUDY, PHASE II

Subjects Tested

A total of 138 fourth and sixth grade subjects were tested during Phase II of the study. Originally 151 subjects had been chosen but several of these were not available for testing. Table 25, page 85 shows the distribution of these subjects according to grade and the sequence to which they had been assigned. The procedures for identifying the "average" and "musically gifted" subjects have been described in Section IV of this report. Certain additional information was obtained for each subject and included: (1) age, in months, through March 1, 1959; (2) I.Q., as determined by the California Test of Mental Maturity, Form S for Elementary Grades; and (3) the kind of musical instrument played and the length of time it had been studied.

Table 26, page 85, summarizes the information regarding age and I.Q. The figures given are the means for each group in each school. This shows relatively little variation in age for each group in the three schools. There is a greater difference between the means of the various groups in terms of I.Q. The "gifted" group in each grade has a higher mean I.Q. than the "average" group in the same grade and this difference is maintained for each of the three schools. It is of interest to note that the mean I.Q. of subjects in School C is higher than in either of the other two schools, with School A slightly lower than School B. Unfortunately it has not been possible, within the scope of this study, to ascertain either the significance of the differences or the possible effect I.Q. might have upon the subject's

Table 25

Table Showing the Distribution of the 138 Subjects Tested in Phase II

School	4th Grade				6th Grade				Total
	Sequence "A"		Sequence "B"		Sequence "A"		Sequence "B"		
	Average	Gifted	Average	Gifted	Average	Gifted	Average	Gifted	
A	7	6	8	4	14	9	9	6	63
B	4	4	5	3	4	5	5	3	33
C	5	6	6	4	4	8	4	5	42
Totals	16	16	19	11	22	22	18	14	138

Table 26

Table Summarizing the Ages and I.Q. of the 138 Subjects in Phase II.

School	AGE				I.Q.			
	4		6		4		6	
	Av.	Gifted	Av.	Gifted	Av.	Gifted	Av.	Gifted
A	118.9	119.1	141.5	141.7	107.9	111.5	110.1	114.1
B	116.4	116.6	144.9	140.6	114.2	118.0	111.2	120.1
C	117.4	114.9	142.0	142.0	120.3	127.6	116.3	125.1
CS*	117.8	116.9	142.5	141.6	115.2	119.2	111.8	118.5
CAG**	117.4		142.0		115.8		115.2	

* "CS" - Combined Schools

performance in the tests of Phase II. Limited research on the topic tends to show that a high level of native musical capacity is not necessarily accompanied by a similar superiority of intelligence. There is little information available regarding the possible effect I.Q. might have upon musical achievement although it may be assumed that these two are more closely related than I.Q. and musical capacity. The extent of this latter relationship needs to be determined by careful research.

There is reason to believe that experience in playing a musical instrument results in a higher level of music reading ability because instrumental performance tends to crystallize the concepts of pitch and tonal relationships as expressed by musical notation. This information was secured for each subject participating in Phase II and is summarized in Table 27. The data includes only those subjects reporting instrumental training of 12 months or more. The writer assumed that a shorter period of instruction would probably have little effect upon the ability to read music. The table clearly shows that more "gifted" subjects had such training than "average" subjects. The total N of 16 fourth grade subjects accounts for 25.8 per cent of the total fourth grade sample; the 37 sixth grade subjects are 48.7 per cent of the sixth grade sample. When the total group of 138 is considered, the data shows that 53 subjects, or 38.4 per cent had at least 12 months of instruction. The differences between the three schools, in terms of the sample, may be more clearly shown when the data is expressed as "per cent of the total sample", as in Table 28.

On the basis of the data it would appear that the differences, in terms of I.Q. and instrumental training, are greater between the "average" and "gifted" groups than between the schools.

Table 27

Table Summarizing the Mean Number of Months of Instrumental Instruction for 53 Subjects in Phase II

School	4th Grade				6th Grade			
	Average		Gifted		Average		Gifted	
	N	Mean	N	Mean	N	Mean	N	Mean
A	0	0	6	23	4	15	11	32
B	2	24	4	21	3	19	5	35.6
C	1	24	3	44	4	31.5	10	38.4
Totals	3	24	13	27.2	11	22.2	26	35.9
Grand Total	N = 16		Mean = 26.6		N = 37		Mean = 31.8	

Table 28

Table Showing Per Cent of Sample Reporting Instrumental Instruction

School	4th Grade			6th Grade			4th & 6th Grade		
	N	n	%	N	n	%	N	n	%
A	25	6	24.0	38	15	38.5	63	21	33.3
B	16	6	37.5	17	8	47.1	33	14	42.4
C	21	4	19.1	21	14	66.7	42	18	42.9

N - total number of subjects for each school in each grade.
n - number of subjects reporting instrumental instruction.
% - n/N

Test Results, Tonal Configurations

The results of the testing in Phase II will be presented in terms of each task, rather than in terms of each sequence of tasks, in order to simplify the presentation. The two tasks were: (1) to learn five tonal configurations within a 15-minute testing period; and (2) to learn a short song constructed from these five tonal configurations within a similar 15-minute testing period. For Sequence A, the tonal configurations were first, followed in two days by the song; for Sequence B, the tasks were reversed. In order to ascertain which of these two sequences led to more effective learning a third task, common to both sequences, should have been included in the study. Such a task could have been a third song, presumably using the same five tonal configurations, which would have served as a criterion measure. The writer recognized the importance of the third task and such a song was constructed, but it was impossible to secure this additional measure within the time limitations of the present study. A comparison of the two methods, following the procedures of this study, will be done at some subsequent time.

Two kinds of scores will be used to present the results; the "per cent correct" score, and the "rate-of-learning" score. These have been described earlier in the report and it will be remembered that the "per cent correct" score was calculated by first summing the number of correct responses to an item and dividing by the total number of trials for the item; then summing the "per cent correct" scores for each item and dividing by the number of items to obtain the score for the total test.

One of the first steps taken was to ascertain the reliability of the tests used in Phase II. The subjects in each of the two sequences at each

Table 29

Table Showing the "Per Cent Correct" Mean Scores for Two Equal Groups

(A) Tonal Configurations

	4th Grade				6th Grade			
	Sequence A		Sequence B		Sequence A		Sequence B	
	I	II	I	II	I	II	I	II
N	16	16	15	15	22	22	16	16
ΣX	1004	935	1109	1147	1703	1489	1254	1140
M	62.75	58.44	73.93	76.47	77.27	67.68	78.38	71.25
ΣX^2	78190	65945	90505	96113	145175	113671	111270	91666
	t = .4103		t = .2829		t = 1.2719		t = .7217	

(B) Song

	4th Grade				6th Grade			
	Sequence A		Sequence B		Sequence A		Sequence B	
	I	II	I	II	I	II	I	II
N	16	16	15	15	22	22	16	16
ΣX	985	988	547	675	1831	1647	831	855
M	61.56	61.75	36.47	45.0	83.28	74.86	51.94	53.44
ΣX^2	71867	70734	28479	40191	156341	132793	53249	54849
	t = .0203		t = .9123		t = 1.5509		t = .1674	

grade level were divided into two equal groups, assignment to groups being made at random. The "average" and "gifted" pupils were combined, as were the three schools. If the differences between the means for each of these two groups was not significant it could be assumed that the tests were reliable and that any other differences were due to factors other than the tests. Table 29 presents the data for both the learning of tonal configurations and the learning of the song. The "t" values for the split groups in each grade and for each sequence show that no real difference exists between the means and the tests may be considered reliable for this particular sample. It should be pointed out that the scores on the song represent the percentage of correct responses to the tonal configurations within the song.

Analysis of the "per cent correct" scores on tonal configurations presented individually is summarized in Table 30, page 91. The highest possible score is 100, indicating that all five items were correct on every trial taken by the subject. The reader is reminded that for Sequence A this was the initial task but for Sequence B it was the second task and was performed after the subject had had practice in learning the song. The higher means for average and gifted subjects in each grade who were tested in Sequence B are probably due to this practice which they received. Examination of the individual tabulations shows that a total of 12 fourth grade pupils and 20 sixth grade pupils had scores of 100. This represents 23.2 per cent of the total sample of 138. More gifted subjects, 10 fourth grade and 18 sixth grade, had perfect scores and at the sixth grade level 13 of the 20 perfect scores were in Sequence A. The fourth grade showed 9 perfect scores for Sequence B. It would seem that the practice effect for Sequence B is

Table 30

Table Showing the Mean "Per cent Correct" Scores for Average and Gifted Subjects on the Tonal Configurations

	4th Grade				6th Grade			
	Sequence A		Sequence B		Sequence A		Sequence B	
	Average	Gifted	Average	Gifted	Average	Gifted	Average	Gifted
N	16	16	19	11	22	22	18	14
ΣX	602	1337	1224	1032	1227	1985	1203	1275
M	37.63	83.56	64.42	93.82	55.77	90.23	66.83	91.07
ΣX^2	29978	114157	88624	97994	74955	183891	90927	117031

Table 31

Table Showing the Mean "Rate of Learning" Scores for Average and Gifted Subjects on the Tonal onfigurations

	4th Grade				6th Grade			
	Sequence A		Sequence B		Sequence A		Sequence B	
	Average	Gifted	Average	Gifted	Average	Gifted	Average	Gifted
N	16	16	18	9	22	17	18	11
ΣX	97.2	533.4	262.5	601.1	299.5	1016.4	309.4	528.2
M	6.08	33.34	14.58	66.79	13.61	59.79	17.19	48.02
ΣX^2	929.44	35215.4	12505.9	53773.8	12985	82015.8	14265.7	36710.6

more evident at the fourth grade level than the sixth grade level. It is also possible that the gifted sixth grade subjects in Sequence A were more competent than their peers in Sequence B.

The data in Table 30 shows the relative superiority of the gifted subjects as compared with average subjects. For each sequence in each grade the gifted group has a considerably higher mean, a performance which could have been predicted without making the study. Of more interest and of greater significance is the comparison of each group in Sequence A with its parallel group in Sequence B to ascertain the degree to which practice on a related task influences test scores. The "t" tests were:

1. 4th grade, average, Sequence A
2. 4th grade, average, Sequence B
3. 4th grade, gifted, Sequence A
4. 4th grade, gifted, Sequence B
5. 6th grade, average, Sequence A
6. 6th grade, average, Sequence B
7. 6th grade, gifted, Sequence B
8. 6th grade, gifted, Sequence B

t12= 3.4689, there is a significant difference between the means at the .01 per cent level.

t34= 1.9550, there is no significant difference between the means.

t56= 1.6422, there is no significant difference between the means.

t78= .1897, there is **no** significant difference between the means.

These data appear to support the following conclusions:

1. Practice is more necessary for average fourth grade pupils than any other group. This group needs to develop an understanding of the concepts involved in reading musical notation.

2. Gifted pupils in both grades already appear to understand these concepts and derive comparatively little benefit from practice since

their scores are already relatively high.

3. Average sixth grade pupils do not show significant improvement as a result of practice but this may be due to a number of factors such as a general lack of interest in reading music; not fully understanding the concepts involved in music reading; and lack of experience in reading independently. This study does not attempt to identify the the factors which may be present.

The writer was particularily interested in attempting to identify the rate at which pupils in these two groups were able to learn the five test items. For this purpose a "rate-of-learning" score, described in Section V, was used as a means of identification. This score is based on the gains made by the subject, these gains being divided by the number of trials taken by the subject as an indication of the rate at which the gains were made. The initial trial was not included in the total number of trials. This score did have several disadvantages but the writer was not able to find a scoring system which could be adapted to the present study. The disadvantages are as follows:

1. The subject earning a perfect score on the initial trial cannot improve his score and consequently receives an RL score of 0. In treating the data the writer arbitrarily excluded those subjects for whom no rate of learning would be possible.

2. The subject who either made no gains from the initial trial to the final trial or who earned a lower score on the final trial than on the initial trial, received an RL score of 0. These subjects are included in the data because the opportunity for learning was always present. The reason for assigning a 0 to the subject whose score be-

came poorer was a device to avoid the problems inherent in working with negative values.

3. The RL score is based on the total test and is not, as is the case with the mean "per cent correct" score, an average of the performance on each item. This works to the disadvantage of the subject who requires many trials to learn a single item, the other items being correct within the first few trials. In view of the totality of the task which involved learning the song, the writer decided to score the tonal configurations as a group rather than individually. One reason for the difficulty encountered in scoring the tasks in this study is due to the complex nature of the tasks themselves. Research activity in the field of music although calling attention to the existence of a scoring problem, fails to recommend more adequate methods of scoring.

The highest RL score a subject could earn is 100, possible only if the subject needed only a second trial on the test to learn the items not responded to correctly on the initial trial. There were 9 fourth grade subjects and 12 sixth grade subjects with RL scores of 100. An additional 3 fourth grade subjects and 8 sixth grade subjects did not have an RL score because they made correct responses to all the items on the first trial of the test. Table 31, page 91 summarized the RL scores for the sample. The difference between the two groups in each grade is quite evident, with the mean for gifted subjects being four to five times as large as that for average subjects. The table also shows that the means for subjects in Sequence B were higher than the means for the parallel group in Sequence A, with the exception of the gifted pupils in sixth grade.

To ascertain the degree to which practice on a related task influenced the RL scores of subjects in Sequence B, "t" tests for the significance of the differences in the means were made. The "t" tests were:

1. 4th, average, Sequence A
2. 4th, average, Sequence B
3. 4th, gifted, Sequence A
4. 4th, gifted, Sequence B
5. 6th, average, Sequence A
6. 6th, average, Sequence B
7. 6th, gifted, Sequence A
8. 6th, gifted, Sequence B

t_{12} = 1.8036, there is no significant difference between the means.

t_{34} = 2.1846, there is a significant difference between the means at the 5% level.

t_{56} = .5356, there is no significant difference between the means.

t_{78} = .8592, there is no significant difference between the means.

These data parallel the results of the "t" tests made to test the significance of the differences between the means of the "per cent correct" scores, a result which was predictable because of the similarity of the two kinds of scores.

The reader will recall that for Phase I the differences in mean scores for fourth and sixth grade had given a t of 1.8397 and it was concluded that the differences were not significant. The 138 subjects in the present sample were considered only in terms of grade; no differentiation being made between average and gifted or between the two sequences. The t test gave a t of .9169 and it was again concluded that the difference between the two grades, in terms of mean Rl scores, was not significant. The mean for the fourth grade was 26.17; for the sixth grade, 31.67.

It will be remembered that in Phase I the scores for the three schools showed a significant difference existing between School B and the other two schools but that the difference between Schools A and C was not significant. The sample for Phase I was a random sample while the subjects for Phase II represented two distinct levels of competence, with only the average group selected on a random basis. To ascertain whether School B continued to rank the lowest in terms of achievement the RL scores on the tonal configurations for the three schools, combining grades and groups, were analyzed. Analysis of variance showed that there was a significant difference between the three schools at the 5% level, F= 4.5328. The data is presented in Table 32. The t tests between the individual schools showed the following:

TABLE 32

Analysis of Variance Between Schools
for the Tonal Configurations of Phase II

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
Total	144347.7997	127	
Schools	9760.9410	2	4880.4705
Pupils in Schools	134586.8587	125	1076.6949
			(F= 4.5328)

1. Between Schools A and B, t= 1.3265, and the difference is not significant.
2. Between Schools A and C, t= 2.9211, and the differences are significant at the one per cent level.
3. Between Schools B and C, t= 1.764, and the differences are not significant.

The means of these three schools were: School A, 21.47; School B, 26.99; and School C, 42.09. The data tends to support the following conclusions:

1. Subjects in School C performed at a high level of competence in both Phases I and II.

2. Subjects in School A performed similarily to those in School C for Phase I but on Phase II their performance more closely paralleled that of School B.

3. Subjects in School B ranked lowest in performance on Phase I but paralleled School A on Phase II.

4. The change in position of School A for the two phases may be due to the differences in samples. It is also possible that Phase II represents a more accurate evaluation of the pupils in the schools than did Phase I. If this is the case, then it may be concluded that subjects in Schools A and B are similar while the subjects in School C are different, in terms of general competence in reading music.

Test Results, Song

For Sequence A the song followed a 15-minute learning period of the five tonal configurations; for Sequence B it was the initial task. Analysis of the "per cent correct" scores, with 100 the highest possible score, showed a total of only 10 perfect scores. It will be remembered that in scoring the song, the score unit was the tonal configuration within the song. Two gifted fourth grade subjects, one in each sequence, and 8 gifted sixth grade subjects in Sequence A account for the perfect scores. This is a partial indication that learning a song is a more difficult task than learning individual tonal configurations, for which 32 perfect scores were obtained. The data in

Table 33, page 99, again indicates the superiority of the gifted subjects. Comparisons between the groups for the two sequences shows the following:

1. For the 4th grade, average, Sequence A with Sequence B, t= [illegible], indicating that the difference between the means is significant at the one per cent level.

2. For the 4th grade, gifted, Sequence A with B, t= 2.0170, indicating no significant difference between the means.

3. For the 6th grade, average, Sequence A with B, t= 5.2728, indicating a significant difference at the .001 level.

4. For the 6th grade, gifted, Sequence A with B, t= 4.5811, indicating a significant difference between the means at the .001 level.

These results support the following conclusions regarding the importance of practice in learning to read a song:

1. Only the gifted fourth grade subjects did not appear to be influenced by the presence or absence of practice on tonal configurations prior to learning a song; this is, of course, true only for this particular sample.

2. For all other groups, especially at the sixth grade level, prior practice on tonal configurations resulted in significant improvement in reading the song. This was not evident when the scores on tonal configurations were compared on the basis of prior practice on the song. This tends to indicate that the reading of songs should be preceded by some prior related activity, not necessarily learning tonal configurations, so the subject moves from the simple to the complex and is aware of what is taking place.

Table 33

Table Showing the Mean "Per cent Correct" Scores for Average and Gifted Subjects on the Song

	4th Grade				6th Grade			
	Sequence A		Sequence B		Sequence A		Sequence B	
	Average	Gifted	Average	Gifted	Average	Gifted	Average	Gifted
N	16	16	19	11	22	22	18	14
ΣX	690	1283	484	738	1466	2012	667	1019
M	43.13	80.19	25.47	67.09	66.64	91.45	37.06	72.79
ΣX^2	36572	106029	15432	53238	103322	185812	30909	77189

Table 34

Table Showing the Mean "Rate of Learning" Scores for Average and Gifted Subjects on the Song

	4th Grade				6th Grade			
	Sequence A		Sequence B		Sequence A		Sequence B	
	Average	Gifted	Average	Gifted	Average	Gifted	Average	Gifted
N	16	16	19	11	22	20	18	14
ΣX	91.7	314.4	81.3	236.1	277.4	966.0	99.3	302.6
M	5.73	19.65	4.28	21.46	12.61	48.3	5.52	21.61
ΣX^2	836.99	14006	565.5	12892.8	4520.9	70862.7	957.55	9192.1

The results agree, in general, with the t tests made for the "per cent correct" scores on the song and would indicate that the more musically mature and experienced sixth grade subject is better able to profit from certain types of related practice than the fourth grade subject.

The four groups in each grade were combined into a single group and the RL scores for the two grades compared. The data showed a mean of 13.15 for fourth grade; 21.19 for sixth grade. The t test gave a t of 2.0532 and the difference between the means of the two grades was considered to be significant at the 5% level. This would also support the conclusion that the maturity of the sixth grade subject gives him an advantage when the rate of learning is considered.

It has been established that for Phase I and for the tonal configurations of Phase II, the differences between the three schools was significant. The means of the RL scores on the song, combining all groups within each school, were as follows: School A, 15.09; School B, 17.48; and School C, 19.10. Analysis of variance showed that these differences were not significant at the 5% level, F= .4182.

TABLE 35

Analysis of Variance between Schools for the Song of Phase II

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square
Total	65770.5199	135	
Schools	411.0686	2	205.5343
Pupils in Schools	65359.4513	133	491.4244
			(F= .4182)

PRECEDING PAGE MISSING

It may be concluded that the samples from the three schools vary considerably when comparisons are made in terms of achievement scores, such scores reflecting previous musical training and experience; but that these same samples do not vary significantly when comparisons are made in terms of the ability of the subjects to learn on the song. The tasks which served as a basis for comparing schools in Phase I and the tonal configurations of Phase II evidently gave the subjects a greater opportunity to utilize previous experiences than did the learning of the song. This also supports the generally accepted point-of view that reading a song is a highly complex process.

Test Results, Identity of Tonal Configurations

One of the major purposes of this exploratory study was to test the hypothesis that tonal configurations which are readily perceived in isolation tend to retain their identity when presented in the usual context of a song. The data used for this aspect of the study is the "per cent correct" score for each tonal configuration in the five-item test and the same score for each item in the song. The product-moment correlations for all groups on each item were calculated and the basis for comparison is the size of the resultant "r". The raw data will not be used in this section of the report in order to simplify and condense the presentation. For the same reason, no attempt is made to retain the identity of the individual schools.

The mean "per cent correct" scores for all items in the test were compared to the same scores for the song, giving the following correlations for the combined schools:

1. 4th, average, Sequence A r= .32
2. 4th, average, Sequence B r= .71

3. 4th, gifted, Sequence A r= .42
4. 4th, gifted, Sequence B r= .44
5. 6th, average, Sequence A r= .58
6. 6th, average, Sequence B r= .37
7. 6th, gifted, Sequence A r= .41
8. 6th, gifted, Sequence B r= .60
9. 4th, average and gifted, Sequence A r= .72
10. 4th, average and gifted, Sequence B r= .74
11. 6th, average and gifted, Sequence A r= .76
12. 6th, average and gifted, Sequence B r= .61

When the average and gifted groups are considered separately, the correlations are higher for groups in Sequence B than those in Sequence A except for the average group in sixth grade, with the difference between the two sequences for the gifted fourth grade subjects being relatively insignificant. The opposite holds true for the average sixth grade subjects. When the two ability groups are combined, the differences between sequences for fourth grade become less evident and the sixth grade shows a larger correlation for Sequence A. It is to be noted that these correlations are higher than those for the individual groups. For Sequence A, 4th grade average was .32 and 4th grade gifted was .42 but when the two groups are combined the "r" becomes a .72. It is the opinion of the writer that these higher correlations result from the fact that the average subjects are at the lower end of the distribution and the gifted subjects at the upper end, thus yielding a higher correlation than when each group is treated separately.

These preliminary correlations show that there is a tendency for tonal configurations to retain their identity. In order to present more complete data the same kinds of correlations were calculated for each of the five tonal configurations used in Phase II. These will be discussed individually and then summarized at the conclusion of the section. In preference to a summary table, the means for each item, together with the correlations, will be presented at the time each item is considered. Mx will designate the mean of the item as presented in tonal configurations, regardless of sequence, and My will refer to the mean of the same item as it appeared in the song.

Item 1: ($\underline{S}\ \bar{D}\ \bar{M}$)

	Sequence A				Sequence B			
	Average		Gifted		Average		Gifted	
	Mx	My	Mx	My	Mx	My	Mx	My
4th	23	66	71	89	74	64	94	87
	r =	.10	r =	.08	r =	-.56	r =	.44
6th	49	87	78	99	72	61	88	89
	r =	.45	r =	-.16	r =	.37	r =	-.04

This item was the first measure of the song and one would anticipate, because of its unique character, that subjects would experience little difficulty in recognition. The means of the various groups show that sixth grade subjects are more familiar with the item than are fourth grade subjects, the exception being the average sixth grade group in Sequence B. There is considerable variation for the correlations and only three of these appear to show that subjects recognized the item when it appeared in a different context than its initial presentation. For Sequence A the average sixth group seemed to be most aware of the item; for Sequence B the average sixth and gifted fourth groups are consistent. The high negative correlation for the average fourth group, Sequence B, indicates considerable shifting of position from task to task. This shifting is not unusual but it does present problems of interpretation. The remaining four correlations are so low as to imply that little identification took place.

Item 2: (M R D)

Treatment I

	Sequence A				Sequence B			
	Average		Gifted		Average		Gifted	
	Mx	My	Mx	My	Mx	My	Mx	My
4th	62	64	91	98	76	44	83	76
	r= .10		r= .93		r= .02		r= .28	
6th	69	85	96	98	75	53	92	92
	r= .20		r=.29		r=-.03		r=-.14	

Treatment II

	Mx	Mz	Mx	Mz	Mx	Mz	Mx	Mz
4th	62	33	91	76	76	15	83	67
	r= -.17		r= .38		r= .41		r= .53	
6th	69	58	96	90	75	19	92	67
	r= .32		r= .08		r= .34		r= .02	

Treatment III

	My	Mz	My	Mz	My	Mz	My	Mz
4th	64	33	98	76	44	15	82	67
	r= .79		r= .52		r= .48		r= .93	
6th	85	58	98	90	53	19	92	67
	r= .22		r= .28		r= .30		r= .26	

This item is the simplest [illegible] which occurred twice within the song. Treatment I is the comparison of the individual item with the first appearance of the same item in the song. Treatment II compares the individual item with the second appearance of the item in the song and treatment III shows the relationship of the two items as they appear within the song.

All groups showed relatively high means on the item, the lowest means are on the second appearance of the item in the song for the average sixth group in Sequence B. In general, the individual item shows higher means than the item within the song. For treatment I, the only group to show consistent recognition is the gifted fourth, Sequence A. The data showed that the first appearance of the item in the song tended to conceal its true identity. The means for the second appearance in the song, treatment II, are uniformly lower for each group than the means for the first appearance. There is a slightly higher amount of recognition for the second appearance of the item as evidenced by the five "r's" above .30. Treatment III shows that two appearances of the same item within a song tend to increase the degree of recognition which takes place.

All "r's" are positive and five are .30 or higher. No single group shows a consistent performance on the item, the higher "r's" shifting from group to group for each treatment. The lack of stability suggests that subjects are not aware that the item exists as a note-grouping, but are reading it as unrelated individual notes. The two appearances within the song emphasize the grouping because the item has the same rhythmic structure in both instances and is a complete measure, the measure bars setting it apart as an entity.

The data also shows that the "r's" in treatments I and III are higher for Sequence A than for Sequence B; while in treatment II, Sequence B is the preferred sequence. This suggests that for this item those subjects having prior practice on the individual items carried over more recognition than the subjects who began with the song and related this experience to the individual items. The data also shows that for Sequence B the fourth grade is more apt to recognize the item in a different context than is the sixth grade. The additional experience may cause the sixth grade subject to read less carefully and "invent" a response which seems more appropriate, in terms of the total situation, than the item actually presented to him.

Item 3: (S M F R)

	Sequence A				Sequence B			
	Average		Gifted		Average		Gifted	
	Mx	My	Mx	My	Mx	My	Mx	My
4th	28	37	75	78	40	8	80	42
	r=.44		4= .40		r=.70		r=.64	
6th	45	61	93	95	58	23	92	64
	r= .20		4=-.16		r= .39		r= .73	

Although only the gifted fourth in Sequence A show high mean scores on the item, all but one of the "r's is positive, with six being .39 or higher. The disjunct, but sequential characteristic of this item appear to simplify the recognition process since subjects remember the uniqueness of the item even though they may not always be able to sing it correctly. The fourth grade subjects show a much higher degree of recognition and greater consistency than do the sixth grade subjects, although the latter earned higher means throughout.

For this item the subjects in Sequence B show a higher degree of recognition than subjects in Sequence A which suggests that the appearance within a song is more meaningful than an isolated presentation.

Item 4: (F M R D)

	Sequence A				Sequence B			
	Average		Gifted		Average		Gifted	
	Mx	My	Mx	My	Mx	My	Mx	My
4th	51	22	91	54	78	10	97	47
	r= .25		r= .18		r= .35		r= .14	
6th	64	44	90	75	56	24	93	52
	r= .17		r= .65		r= .15		r= .30	

Item 4 proved to be more difficult in the song than as an individual item and this is shown by the lower My values. The chief reason for this was ecause it began the interval of a fourth above the final tone of he previous measure and subjects experienced difficulty in locating the correct tone. Although only three of the correlations are .30 or higher, all "r's" are positive, indicating recognition of the item by many subjects. The size of the "r's" for the various groups does not show a consistent pattern, nor does one sequence seem to be preferred to the other although the gifted subjects in Sequence A do better than their peers in Sequence B.

Item 5: (S L S F M)

	Sequence A				Sequence B			
	Average		Gifted		Average		Gifted	
	Mx	My	Mx	My	Mx	My	Mx	My
4th	19	31	84	86	54	14	100	70
	r= .50		r= .34		r= .44		r= 0	
6th	51	61	95	97	73	42	91	81
	r= .39		r= .13		r= .49		r= .72	

The means of the gifted subjects in both sequences continues to exceed the means of the average subjects, a behavior that has been consistent throughout the testing. The identity of the item appears to be retained since all "r's" are positive. For Sequence A the fourth grade subjects, with higher "r's" than sixth grade subjects, apparently are more aware of the individual character of the item even though their mean scores are lower. In Sequence B the sixth grade has the higher "r's". The data shows that subjects within most groups shift their position from task to task, a shifting that has been evident in the analysis of each item.

Summary of the Findings

1. The average fourth grade and gifted sixth grade subjects in Sequence B show a higher degree of recognition of tonal configurations than do these same groups in Sequence A when the correlations between items on the song and individual items is considered. This suggests that subjects identify tonal configurations presented individually more readily if they have had previous opportunity to observe these configurations in a song. It also implies that for these same groups, practice on tonal configurations does not lead to greater recognition of these same configurations in the song. The reverse holds for the average sixth grade subjects in the sample.

2. The subjects within all groups do not perform consistently, as evidenced by the variation in the size of the "r's" for each item. The factors which cause this shift cannot be identified at the present time but it is reasonable to assume, in view of the reliatbility of the test, that the effort and attitude of the subject are significant factors.

3. The data from the present study does not offer strong support for the point-of-view that tonal configurations tend to retain their identity in different contexts unless one is willing to accept the coefficient of correlation obtained for the combined gifted and average subjects in each sequence as sufficient evidence of such retention. These were all relatively

high, .61 to .76, and may represent a more dependable value than the "r's" obtained for each of the sub-groups.

4. The item, if it is to retain its identity from one situation to another, must first be learned. The testing carried on in Phase II did not result in such learning for the majority of the subjects. Shorter, more numerous practice sessions would eventually result in learning the items and a second task presented at that time might have significantly raised the "r's" for the groups.

5. The variability of the "r's" makes it impossible to identify any one sequence of tasks as the most desirable for the development of item recognition. The addition of a third task might well have provided the kinds of data needed to make such a judgment.

6. For all items, the gifted subjects in each grade for each sequence had higher scores than did the average subjects. In 20 of the 28 comparisons of scores the sixth grade gifted subjects scored higher than the fourth grade gifted. This is not the same when the "size" of the "r's" are compared. For these same groups the data shows that in Sequence A the fourth grade gifted group had higher "r's" than the sixth grade gifted group in five out of seven such comparisons; for Sequence B the fourth grade gifted group was higher than the sixth grade gifted in four out of seven comparisons.

7. The same results appear when the average groups in each grade are compared. The sixth grade scores are higher in each case than fourth grade scores while the "r's" for fourth grade groups are higher in nine of the fourteen comparisons than the "r's" of the sixth grade groups.

8. When each item is considered individually no single item shows a uni-

formly high degree of recognition by all the various groups. Item 3 and the two appearances of Item 2 within the song appear to retain this identity better than any other item. The data shows that for these items the fourth grade "r's" are higher than sixth grade in 13 out of 16 comparisons. For Items 1, 4 and 5 the sixth grade "r's" are higher than those of the fourth grade in 8 of 12 comparisons. This is a more reliable indication of what is taking place and shows that fourth grade subjects tend to retain items that are either very simple or which have distinctive characteristics. When the item is neither of these the sixth grade subjects are better able to detect the subtle differences and identify the item in a new context.

9. It may be concluded that the sixth grade subjects in Phase II of the study not only demonstrated a higher level of achievement but also showed a greater degree of item recognition for the test as a whole.

The Influence of Instrumental Instruction on Test Scores

It has been mentioned earlier that many music educators feel that experience in playing a musical instrument results in a higher level of music reading ability because instrumental performance tends to clarify the abstractions of musical notation. A total of 53 subjects in Phase II had reported instrumental training of at least 12 months or more. The "mean per cent correct" scores on the tonal configurations were used as the basis of comparing the pupils within groups who had training with those having no such training. The results are summarized as follows:

Group	Without Training		With Training		"t"
	N	Mean	N	Mean	
4th Average, Sequence A	15	39.3	1	13	
4th Average, Sequence B	16	66.8	3	85.3	1.3936
4th Gifted, Sequence A	4	79.3	12	85.0	.7763
4th Gifted, Sequence B	6	90.3	5	98.0	1.1936
6th Average, Sequence A	17	54	5	57.6	.3339
6th Average, Sequence B	12	66	6	68.5	.1952
6th Gifted, Sequence A	6	84.0	16	92.5	1.1961
6th Gifted, Sequence B	5	91.8	9	90.7	.2330

The "t" tests made show that the differences between the means were not considered significant at the 5% level and it may be concluded that the instrumental training had little effect upon the test scores of the subjects in the sample.

There are several factors which need to be considered before any conclusions can be reached:

1. The groups compared are all relatively small.
2. Piano was the instrument mentioned most frequently and this may not influence music reading ability, in the sense of this study, as much as a single-line instrument like the violin.
3. The nature of the instruction may have given more emphasis to technical performance than to developing skill in reading music.
4. The instrument may not have been studied long enough to enable the child to develop an adequate concept of the process of music reading.
5. The tests may not really give a true measure of music reading ability, especially since the use of a vocal response may have served

as a deterent.

For these reasons the writer is not willing to go beyond the limitations of the present study.

Relationship Between Scores on Tonal Configurations and the Teacher Ratings and Kwalwasser Tests

The writer has already discussed the problems associated with the identification of the musically gifted children. There are many possible definitions of gifted children and this study defined the musically gifted subject as the child whose musical achievement was superior to that demonstrated by the majority of his classmates. This superior achievement results from favorable attitudes toward musical activities, general knowledge about music, and performing skill in singing and playing instruments. The experienced music teacher, having contact with the children over an extended period of time, should be able to evaluate those outstanding children in each group. The Kwalwasser Music Test does not measure achievement but seeks to identify native capacity.

Data presented earlier in the report showed the correlations between the teacher's ratings and the Kwalwasser Test to be relatively low for all schools. This is to be expected because the two criteria do not measure exactly the same things. To ascertain which criterion measure, if any, had the closest relationship to the reading of tonal configurations in Phase II, the product-moment correlations were computed for the average and gifted subjects in each school. These correlations are represented as AB and AC:

1. A represents the scores on tonal configurations.
2. B represents the scores on the teacher ratings.
3. C represents the scores on the Kwalwasser test.

These "r's" are given for each grade in each school as follows:

School	Grade	Average or Gifted	rAB	rAB
1. A	4	A	.60	-.06
2. A	4	G	.21	-.10
3. A	6	A	.08	-.14
4. A	6	G	.48	-.36
5. B	4	A	.79	.17
6. B	4	G	-.03	.59
7. B	6	A	-.28	-.07
8. B	6	G	-.29	.17
9. C	4	A	-.25	.42
10. C	4	G	.03	.17
11. C	6	A	.13	-.69
12. C	6	G	-.10	-.09

This data shows a lack of consistency, not only from school to school but for the various groups as well. The data shows the following:

1. Neither of the criterion measures are able to show a consistent relationship with the task used in Phase II although the teacher's rating correlates higher than the Kwalwasser for all groups in School A.
2. For School B the preference is given to the Kwalwasser test although the correlations are so low as to indicate that such a basis is not reliable. School C is similar to School B.
3. When the gifted subjects are considered, the higher "r's" occur for the Kwalwasser in four out of six such comparisons.
4. When the average subjects are considered, the higher "r's" occur for the teacher's ratings.

On the basis of the data obtained the writer finds himself in agreement with the many other investigators who have attempted to establish some basis for identifying musically gifted individuals, that this basis is yet to be discovered. It is possible that the most adequate basis for making such identification is the performance on the tasks in Phase II rather than outside criteria. It is also quite possible that the absence of a large number of significantly high "r's" and the lack of a consistent pattern shows that teachers are not in a position to adequately evaluate the music reading ability of their pupils because there is comparatively little opportunity for individual attention within the classroom situation.

SUMMARY AND CONCLUSIONS

Summary of Major Findings Based on the Analysis of Data

1. There is no significant difference between boys and girls in terms of the ability to read music.

2. For Phase I there were no significant differences between the scores earned by fourth, fifth and sixth grade subjects. For Phase II there were no significant differences between fourth and sixth grades either in terms of general scores or in terms of the rate of learning scores for tonal configurations. However, the RL scores on the song showed a significant difference between these two grades.

3. When "per cent correct" scores were used there were significant differences between the three schools; School C, while comparable to School A in Phase I, was higher than the other two Schools in Phase II. School B ranked lower on Phase I and was comparable to School A on Phase II. There were no differences when RL scores on the song were used as a basis for comparison.

4. The analysis of data for Phase I showed:

    a. The subjects experienced considerable difficulty in reading tonal configurations which commonly appear in the songs they sing.

    b. The common reading errors showed a general tendency to respond to the outward shape of the configuration but to substitute a similar, more familiar configuration for the unfamiliar one being read.

3

c. Considerable guessing takes place when subjects attempted to read, as evidenced by the number of different errors which were made.

d. The subjects performed at a much higher level on the aural part of the test than on the visual part. There were few common errors, the major cause for error arising out of the subjects apparent inability to retain the sound of the starting pitch throughout the presentation of the item.

e. The relationship between the responses to the aural and visual presentations of an item was not as high as expected, "r's" of .50 to .70. This is, however, not as accurate a comparison as might have been made.

5. The teacher's ratings of subjects were compared with the results of the Kwalwasser test and showed very low correlations.

6. Gifted subjects in Phase II consistently did better than average subjects on each task. They knew more about reading music, as evidenced by the "per cent correct" scores, and learned the material at a rate three to four times faster than average subjects.

7. The song was a more complex task than the individual tonal configurations, regardless of the sequence of tasks.

8. Comparisons between sequences for the various groups, using the scores earned on tonal configurations as a basis for comparison, showed the following:

a. The "per cent correct" scores for groups in Sequence B, with prior practice on the song, were not significantly higher as a result of this practice than the scores for groups in Sequence A who did not have practice. Only the fourth grade average group seemed to benefit from such practice.

b. The RL scores for groups in Sequence B were not significantly higher than the same scores for groups in Sequence A except for the gifted fourth grade subjects.

9. Comparisons between sequences for the various groups, using the scores earned on the song showed the following:

a. The "per cent correct" scores for groups in Sequence A, with prior practice on tonal configurations, were significantly higher as a result of this practice than the scores for groups in Sequence B who had no practice. Only the gifted fourth grade group seemed not to benefit from such practice.

b. The RL scores for sixth grade groups in Sequence A were significantly higher than the scores for the sixth grade groups in Sequence B. There were, however, no significant differences between fourth grade groups in the two sequences.

10. Analysis of the degree to which tonal configurations retain their identity from one situation to another showed that:

a. Average fourth grade and gifted sixth grade subjects show greater recognition when the sequence from song to configuration is followed.

b. Average sixth grade subjects show greater recognition when Sequence A is followed.

c. Considerable shifting of the position of the subjects from one task to another takes place within all groups. This variability makes it difficult to accurately interpret the "r's".

d. Items 2 and 3 show fairly consistent recognition by fourth grade groups; Items 1, 4 and 5 are more readily recognized by by sixth grade groups.

11. The scores of subjects in this study who reported instrumental training were not significantly higher than the scores of subjects who did not have such training.

12. The identification of gifted and average subjects on the basis of teacher's ratings and the Kwalwasser Test does not result in a high level of reliability in terms of this study.

Implications of the Study

1. The time allowed for the tasks in Phase II was not sufficient to result in complete learning by most of the subjects. This suggests that more learning sessions, rather than longer ones, would result in greater understanding of the concepts of musical notation.

2. Accurate reading of musical notation is more likely to result when greater emphasis is placed upon understanding the significance of the notation rather than upon mere "imitation of musical materials.

3. The musically superior child needs to be given more opportunities to develop music reading competence. The data shows both fourth and sixth grade superior children to learn at a much more rapid rate than average children, a fact that must be considered in planning the music activities for the grades.

4. The data did not clearly establish that one sequence of tasks is superior to the other in developing music reading competence. It did show, however, that although prior practice on the song did not appear to have much effect upon the scores obtained for learning individual configurations, learning the song was made more effective through prior

practice on tonal configurations. The fact that the song is usually approached as a complete entity, as it should be, makes it difficult for children to then recognize the individual elements of the song when they are isolated.

5. If the music reading is approached by means of the configuration, greater emphasis needs to be placed upon helping children recognize the shape and design of such configurations. Subjects in this study were unable to identify the configurations in the song which they had previously learned as individual items until they were asked to make such an identification.

6. The results of the study showed that these sixth grade children did not perform significantly better than the fourth grade children although they clearly demonstrated a faster learning rate for the song. This implies that very little actual growth in music reading skill takes place beyond the fourth grade level because the children are not given the necessary musical tools with which to work. The tools in this sense of the work are the knowledges and understandings about musical notation which are developed through independence of action.

7. Instrumental instruction did not seem to be a factor in the music reading abilities of either the average or gifted children in this study. The training will exert a greater influence if greater emphasis is given to the meaning of the musical notation rather than to mere mechanical responses to visual stimuli. It is highly probable that instrumental training at this level is more concerned with developing control over the instrument **so that he is not able to give his attention to other problems of performance.**

8. The procedures followed in this study are adequate but improvements need to be made in methods of scoring and evaluating the data.

9. The problem of the visual perception of music symbols needs tobe extended to ascertaining whether tonal configurations learned in one key are readily perceived when presented in other tonalities where the notation is different.

10. The relationship between auditory and visual perception needs to be more clearly established. Music educators recognize the interdependence of the two but the degree to which one depends upon the other needs to be identified. Many reading errors made in this study were caused by the subject's inability to "hear internally" what the stimulus should sound like and consequently were unable to control the voice.

Concluding Statement

This exploratory study has given the writer a great deal of valuable information and experience. Many possible avenues of exploration have been identified; the basis for developing important changes in the collection and analysis of data has been provided; and the writer has gained confidence in his ability to carry forward future projects in the skill area of music reading.

BIBLIOGRAPHY

1. Bean, Kenneth An Experimental Approach to the Reading of Music, Psychological Monographs, Vol.50, No. 6, 1938.

2. Burmeister, C. A. "The Role of Music in General Education", Fifty-seventh Yearbook for the National Society for the Study of Education, Part I, (Chicago, Ill.: The University of Chicago Press, 1958) Ch. IX, pp.215-235.

3. Brooks, Marian and Brown, Harry Music Education in the Elementary School. (Cincinnati: American Book Co., 1946).

4. Christ, W. B. Reading of Rhythm Notation Approached Experimentally According to the Techniques and Principles of Word Reading (Indiana University, Unpublished Doctor's thesis, 1953).

5. Flanagan, John C. "Considerations in the Selection of Test Items and a Short Method of Estimating the Product-Moment Coefficient From Data at the Tails of the Distribution", Journal of Educational Psychology, 30:674-680 (1939).

6. Jacobsen, O. Irving "An Analytical Study of Eye-Movements in Reading Vocal and Instrumental Music", Journal of Musicology, Vol.3. (1941-42) pp. 1-32; 69-100; 133-164; 197-221; 223-226.

7. Morgan, Hazel N. (Editor) Music in American Education: Source Book Number 2, (Chicago: Music Educators National Conference, 1955).

8. Mursell, James L. "Growth Processes in Music Education", Fifty-seventh Yearbook of the National Society for the Study of Education, Part I, (Chicago: The University of Chicago Press, 1958) Ch. VI, pp. 140-162.

9. Mursell, James L. Human Values in Music Education, (New York: Silver Burdett and Co., 1943).

10. Myers, Louise Kifer Teaching Children Music in the Elementary School. (Englewood Cliffs, N.J.: Prentice-Hall, 1956).

11. Ortmann, Otto On The Melodic Relativity of Tones, Psychological Monographs, Vol. 35, No. 1, 1926.

12. Ortmann, Otto "Some Tonal Determinants of Melodic Memory", Journal of Educational Psychology, 24:454-467 (1933).

13. Ortmann, Otto "Span of Vision in Note Reading", Journal of Proceedings of the Thirtieth Annual Meeting of the Music Educators National Conference, 1937. Pp. 88-93.

14. Pitts, Lila Belle The Music Curriculum in a Changing World. (New York: Silver Burdett Co., 1944).

BIBLIOGRAPHY (Cont'd)

15. Smiley, Edra M. A Study of the Musical Configurations, Symbols, Terms and Words Found in Basic Music Texts at the 4th Grade Level, (Indiana University. Unpublished Doctor's Thesis, 1954)

16. Stokes, Charles F. An Experimental Study of Tachistoscopic Training in Reading Music, (University of Cincinnati, Unpublished Doctor's Thesis, 1943).

17. Weaver, H. E. Studies of Ocular Behavior in Music Reading, Psychological Monographs, Vol. 55, No. 1, 1943.

18. Weaver, H. E. and Van Nuys, Kelvin Memory Span and Visual Pauses in Reading Rhythms and Melodies, Psychological Monographs, Vol. 55, No. 2, 1943.

Appendix A

Song-texts Analyzed for Tonal Configurations

1. *The American Singer, Second Edition*
New York: American Book Company, 1955

Book 2	16 songs
Book 3	19 songs
Book 4	19 songs
Book 5	20 songs
Book 6	22 songs
Total	96 songs

2. *Music For Living*
New York: Silver Burdett Company, 1956

Book 2, *Music in Our Town*	15 songs
Book 3, *Music Now and Long Ago*	16 songs
Book 4, *Music Near and Far*	17 songs
Book 5, *Music in Our Country*	17 songs
Book 6, *Music Around the World*	15 songs
Total	80 songs

3. *Our Singing World*
New York: Ginn and Company, 1951

Book 3, *Singing and Rhyming*	17 songs
Book 4, *Singing Every Day*	19 songs
Book 5, *Singing Together*	19 songs
Book 6, *Singing in Harmony*	21 songs
Total	76 songs

4. *A Singing School*
Boston: C. C. Birchard, 1952

Book 3, *Merry Music*	18 songs
Book 4, *We Sing*	18 songs
Book 5, *Our Land of Song*	18 songs
Book 6, *Music Everywhere*	20 songs
Total	74 songs

Appendix B

FUPIL RATING SCALE

_____________________________ (Name) _____________________ (School)

(Instructions: Please check the response which best describes the pupil's ability or attitude erms of his performance on each item.)

A. GENERAL SINGIN ITY

1. The pupil le to match individual tones and intervals:

_____ Alw
_____ Usually
_____ Occasionally
_____ Seldom
_____ Never

2. The pupil is able to sing a phrase of a song he knows:

_____ But never succeeds, even with help from the piano or teacher
_____ Seldom succeeds, even with help
_____ Needs frequent help
_____ With only occasional help
_____ Independently, with no help

3. In the independent (solo) singing of a song learned by the class, the pupil:

_____ Sings the complete song without help from the teacher or piano
_____ Sings the complete song but needs occasional help
_____ Sings the complete song but needs frequent help
_____ Is seldom able to sing the complete song, even with help
_____ Is never able to sing the complete song, even with help

4. In the independent (solo) singing of a song learned by the class, the pupil:

_____ Almost never sings in tune, usually fails to adjust the pitch when helped
_____ Seldom sings in tune, occasionally is able to adjust the pitch with help
_____ Frequently sings out-of-tune, usually adjusts when given help
_____ Occasionally sings out-of-tune but always adjusts when helped
_____ Almost always maintains correct intonation throughout without help

5. In the independent singing of a song learned by the class, the pupil:

_____ Never makes interval errors, never loses the tonality or key-sense
_____ May make occasional interval errors but regains tonality without help
_____ Makes interval errors and needs help to return to proper tonality
_____ Frequently loses the tonality and does not always respond to help
_____ Almost always loses the tonality and usually fails to respond to help

6. The pupil's singing, either alone or with the class, is characterized by an enthusiastic, eager, whole-hearted and controlled response:

_____ Almost never (0-20% of the time)
_____ Occasionally (21-40% of the time)
_____ Frequently (41-60% of the time)
_____ Usually (61-80% of the time)
_____ Always (81-100% of the time)

B. RHYTHMIC RESPONSIVENESS

7. The pupil, when asked to identify the underlying beat or meter of the song by hearing it, does so immediately and accurately:

_____ Always
_____ Usually
_____ Occasionally
_____ Seldom
_____ Never

8. The pupil, when asked to repeat common rhythmic patterns which he hears, does so immediately and accurately:

_____ Never
_____ Seldom
_____ Occasionally
_____ Usually
_____ Always

C. MUSICAL KNOWLEDGES AND SKILLS

9. When participating in appropriate music reading activities, the pupil's response shows that he is:

_____ a leader, capable of much independent music reading.
_____ partially independent, voice always moves in direction of notes and is frequently accurate.
_____ occasionally able to control the direction of the voice by using notes
_____ seldom is able to react to the notation, learns by rote most of the time
_____ seems to be unaware of the meaning of the notation, cannot differentiate one symbol from another.

10. The pupil demonstrates a knowledge of common musical symbols and terms used in his grade that is:

_____ Poor
_____ Fair
_____ Average
_____ Good
_____ Superior

D. GENERAL INTEREST IN MUSICAL ACTIVITIES

11. In terms of the singing activities carried on, the pupil is:

_____ Very interested and responsive
_____ Usually interested and fairly responsive
_____ Passive, does not put forth much effort but does not resist.
_____ Slightly negative, frequently does not participate at all.
_____ Has a strong negative reaction, usually does not participate

12. In terms of the listening and creative activities carried on, the pupil is:

_____ Strongly negative, fails to pay attention or show interest
_____ Slightly negative, frequently does not show interest
_____ Passive, shows neither a negative or positive reaction
_____ Is usually very interested and fairly responsive
_____ Very interested and responsive